The Dissertation Committee for Marcel Nassar
certifies that this is the approved version of the following dissertation:

Graphical Models and Message Passing Receivers for Interference Limited Communication Systems

Committee:

Brian Evans, Supervisor

Gustavo de Veciana

Robert W. Heath, Jr.

Jonathan Pillow

Haris Vikalo

Graphical Models and Message Passing Receivers for Interference Limited Communication Systems

by

Marcel Nassar, B.E.; M.S.E

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2013

Dedicated to the memory of my uncle,

Hassan.

Acknowledgments

I would like to start by expressing my sincere appreciation to my adviser Prof. Brian Evans for his support and trust. His professionalism, commitment, and belief in my abilities have enabled me to pursue my interests not only within my research field but also within the wider fields of academic knowledge. I would also like to thank my committee members, Prof. Gustavo de Veciana, Prof. Robert W. Heath, Prof. Jonathan Pillow, and Prof. Haris Vikalo for their invaluable comments and suggestions. I greatly benefited from the insights I acquired during their excellent courses at UT Austin and the discussions that I was fortunate to have with them. In addition, I would like to express special gratitude to Prof. Philip Schniter (of The Ohio State University). I was lucky to meet Phil during ASILOMAR 2011 while I was looking into the problem of receiver design under impulsive interference (Chapters 5 and 6). Since then, Phil has been like a technical co-advisor for this dissertation providing in-depth feedback for the last two contributions and co-authoring the resulting papers.

This work was not completed in a vacuum. At UT, I was lucky to be surrounded by brilliant colleagues some of whom provided invaluable feedback on my dissertation and others whom I was lucky to collaborate with on many interesting projects and papers. Thank you Mijung Park, Dr. Omar El Ay-

ach, Dr. Kapil Gulati, Jing Lin, Yousof Mortazavi, Karl Nieman, and Marcus DeYoung. In addition, I would like to thank Dr. Anand Dabak, Dr. Il Han Kim, Dr. Eddie Lin, and Dr. Inching Chen for the interesting and intellectually simulating internships at Texas Instruments and Intel.

At surface, a PhD might appear as a purely academic journey whose crowning achievement is the dissertation. While acknowledging the great effort that goes into the dissertation, I believe the dissertation to be of secondary importance. I have always viewed my grad school years as an opportunity to grow as a person (before the "real" world catches up). As a result, I will always cherish and be indebted to my friends that I was lucky to meet in the great city of Austin, Texas. Thank you Amin, Andras, Lenka, Omar, Mijung, Salam, and Samer.

Last, but not least, I would like to thank my parents Helena and Hussein, my brother Radovan, and my sister Mariam for their support throughout my life. Thank you!

Graphical Models and Message Passing Receivers for Interference Limited Communication Systems

Publication No. _____

Marcel Nassar, Ph.D.
The University of Texas at Austin, 2013

Supervisor: Brian Evans

In many modern wireless and wireline communication networks, the interference power from other communication and non-communication devices is increasingly dominating the background noise power, leading to interference limited communication systems.

Conventional communication systems have been designed under the assumption that noise in the system can be modeled as additive white Gaussian noise (AWGN). While appropriate for thermal noise, the AWGN model does not always capture the interference statistics in modern communication systems. Interference from uncoordinated users and sources is particularly harmful to communication performance because it cannot be mitigated by current interference management techniques.

Based on previous statistical-physical models for uncoordinated wireless interference, this dissertation derives similar models for uncoordinated

interference in PLC networks. The dissertation then extends these models for wireless and powerline interference to include temporal dependence among amplitude samples. The extensions are validated with measured data.

The rest of this dissertation utilizes the proposed models to design receivers in interference limited environments. Prior designs generally adopt suboptimal approaches and often ignore the problem of channel estimation which limits their applicability in practical systems. This dissertation uses the graphical model representation of the OFDM system to propose low-complexity message passing OFDM receivers that leverage recent results in soft-input soft-output decoding, approximate message passing, and sparse signal recovery for joint channel/interference estimation and data decoding. The resulting receivers provide huge improvements in communication performance (more than 10dB) over the conventional receivers at a comparable computational complexity. Finally, this dissertation addresses the design of robust receivers that can be deployed in rapidly varying environments where the interference statistics are constantly changing.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Digital communication systems transmit information bits from a transmitter to a receiver through a physical medium. The main impairments to a communication system, whether wireless or wireline, are the effects of multipath propagation through the physical medium and random fluctuations in the received signal due to disturbance from either natural or man-made sources, referred to as *noise* or *interference*. A common approach for communication system design models the multipath effects as a convolutive linear filter that, in the slow-fading scenario, can be characterized by a fixed impulse response over the duration of one codeword; and the effects of the random distortions as an additive noise process. Statistically, the channel coefficients are modeled as independent complex Gaussian random variables, resulting in the well-known "uncorrelated Rayleigh-fading" and "uncorrelated Rician-fading" models [94]. Similarly, the samples of the additive noise process are modeled as independent complex Gaussian random variables giving rise to the "additive white Gaussian noise" (AWGN) model that dominates the communications literature until today [94].

While an appropriate model for the thermal noise in the receiver cir-

cuitry, the AWGN model fails to capture the characteristics of the noise and interference in modern communication systems. Extensive measurement campaigns in various frequency bands up to 4 GHz demonstrate that the additive noise is in fact *impulsive* with amplitudes up to 40 dB above the thermal background noise [13, 56, 74, 83, 84]. Similar studies for both narrowband and broadband powerline communications (PLC) also indicate that the noise is highly impulsive and in some scenarios occurs in bursts [72, 73, 104].

1.1 Interference in Communication Systems

Broadly speaking, interference or noise refers to the disturbance energy, resulting from either natural or man-made sources, that adds to the transmitted signal at receiver and degrades its ability to successfully detect the transmitted information. Throughout this dissertation, I use noise and interference interchangeably. Interference can be roughly classified based on its source as being either communication or non-communication based. In the following, I give a brief overview of each category.

1.1.1 Non-Communication Based Interference

Non-communication based interference typically includes unintentional electromagnetic emissions from sources that could either be on the same platform as the transceiver, called *in-platform* interference, or external to the platform, called *out-of-platform* interference. The in-platform interference includes the well-known *circuit* noise which is typically assumed to be AWGN. In

the following, I discuss these categories of interference by giving some examples of each.

1.1.1.1 In-Platform Interference

The computational platform contains many subsystems, such as various clocks and buses, that generate emissions that interfere with the transceiver present on the same platform [86]. This interference is not only due to the near-field coupling with radiation at the same frequencies as the driving clocks but also includes the harmonics produced by these subsystems. Given that there are no regulations that limit the interference power inside the platforms and the current trend to decrease their form factors, in-platform interference is becoming a prominent limitation for communications performance in modern devices.

1.1.1.2 Out-of-Platform Interference

This interference is caused by external devices operating in the same frequency band as the communication system. This is common for example in the 2.4GHz Industrial, Scientific, and Medical (ISM) band that is used for IEEE 802.11b/g/n WLANs and Bluetooth which are subjected to interference from various non-communications devices such as microwave ovens [13, 74]. Microwave ovens exhibit non-stationary statistics largely deviating from the Gaussian model and powers as high as 50 dBm at 15 m [13] (comparable to the transmit power of an access point (AP) in WLANs). With a typical

usage on the order of minutes, this can lead to serious disruption for real-time streaming applications such as wireless video and presentations in home and office environments.

The impact of out-of-platform interference is even more severe in PLC networks. PLC interference consists of random impulses of varying durations. It is mainly caused by switching transients of various appliances and devices in individual homes and businesses present on the network [29, 104]. Additional interference can also be picked up by the PLC network acting as an antenna for wireless in-band and aliased signals [29]. This interference power density can reach 50dB above the background noise and is considered one of the main impairments to reliable communications in broadband PLC [104].

1.1.2 Communication Based Interference

The physical propagation medium in both the wireless and PLC systems is shared among many transceivers. Without coordination, the transmissions of these transceivers collide creating interference at each others receivers. Typically, sources that occupy the same frequency band as the signal of interest dominate the resulting interference, labeled *co-channel interference* [37]. A weaker interference source results from transmissions occurring in adjacent bands that leak to the band of interest due to nonlinearities in the transmitter circuitry. This interference is labeled *adjacent-channel interference* [37]. Furthermore, mismatch in channel state information (CSI) can, in many scenarios, lead to residual interference for example due to transmitter inability

to perfectly orthogonalize its transmission with respect to other users.

1.1.2.1 Co-Channel Interference

Co-channel interference arises from overlapping transmissions within the same frequency band. Such a scenario includes *multi-user* systems that employ dense-spatial frequency reuse driven by the increasing demand for higher data rates. A unlicensed band is another example where multiple standards, typically uncooperative, compete for reliable communication under power constraints imposed by the Federal Communications Commission (FCC). Examples include WiFi, Bluetooth, and Zigbee all operating in the 2.4 GHz ISM band [19].

1.1.2.2 Adjacent-Channel Interference

Adjacent-channel interference arises from transmissions in neighboring frequency bands. With the increased push toward smaller form factors, integration of multiple wireless transceivers that operate simultaneously in the same platform is becoming common place allowing users to download data via WiFi while placing a call over the cellular network [77].

1.2 Multicarrier Communication Systems

Due to presence of various scatterers and reflectors in many environments, the transmitted signal typically propagates through multiple paths before reaching the receiver. If the time difference between the received copies

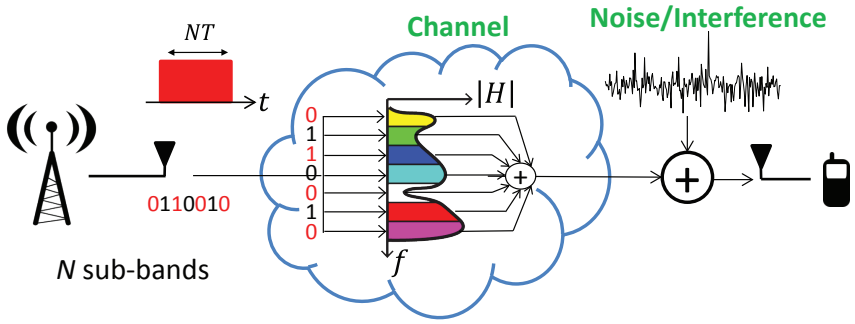


Figure 1.1: Multicarrier Communication: A wide-band channel is divided into a set of narrowband channels with simpler equalization.

of the transmitted signal, called *delay spread*, is significant compared to the signaling time of transmitter, the channel is said to be *frequency-selective*, or *wide-band*. As discussed in the beginning of this chapter, the channel is then modeled as a convolutive linear channel with channel taps given as $\{h_j\}_{j=0}^{L-1}$ with $L > 1$ [94].

In general, having multiple independent copies of the same transmitted signal improves the detection performance through *diversity*; in this case it is *frequency-diversity* [94]. However, if not properly accounted for, wide-band channels can cause significant degradation in communication performance due to the presence *inter-symbol-interference* (ISI), a type of self interference that results from the overlap of the different copies of the transmitted signal arriving at different times at the receiver [94].

Orthogonal frequency division multiplexing (OFDM) is a transmit precoding technique that transforms the wide-band channel into a set of non-interfering, orthogonal, and narrowband sub-channels. This significantly sim-

plifies the channel compensation procedure, called *equalization*, and, under AWGN, leads to independent decoding across the different sub-channels. However, when subjected to interference with typically non-Gaussian statistics, the noise is no longer independent across the sub-channels and independent sub-channel decoding is highly suboptimal [45]. Furthermore, the MMSE channel estimator (important for equalization) is highly non-linear[1] and unknown. By ignoring the statistics of the interference and treating it as an AWGN process, the receiver causes significant degradation (tens of dBs) in communication performance [45]. *This is the central issue addressed by this dissertation.*

1.3 Dissertation Summary

In this dissertation, I address the problem of designing OFDM receivers in interference limited environments. By modeling the non-Gaussian statistics of the interference, I propose a Bayesian inference framework for designing OFDM receivers that jointly estimate the channel/interference and decode the transmitted data. Furthermore, the framework is flexible enough to allow for a trade-off between the communication performance and implementation complexity.

1.3.1 Thesis Statement

In this dissertation, I defend the following thesis statement:

[1]Under AWGN, the MMSE channel estimator is the linear MMSE (LMMSE) [94].

In interference-limited multicarrier communication systems, accurate statistical modeling of the interference enables the design of low-complexity message passing multicarrier receivers that increase the link spectral efficiency by several bits/s/Hz, without any coordination or knowledge of the number, locations, or types of interference sources.

1.3.2 Summary of Contributions

In this dissertation, I investigate the impact of uncoordinated interference on multicarrier systems. I develop statistical models of this interference using (i) physical models of the interference network and stochastic models of interference emissions; and (ii) empirical models based on measurement data collected in the lab and field. I then embed these models as priors in a Bayesian inference framework represented by a factor graph to design message passing receivers that mitigate the effect of the interference and significantly improve the communication performance. Using recent results in approximate inference for high dimensional linear mixing problems [28, 79, 80], the proposed receivers have a low computational complexity on the same order as the typical OFDM receivers in AWGN noise ($\mathcal{O}(N \log N)$ where N is FFT size). The main contributions of this dissertation can be summarized as follows.

1. **Statistical Modeling of Uncoordinated Interference:** In this contribution, I develop a statistical-physical models of uncoordinated interference in PLC networks. In particular, I consider interference emissions that arrive according to a Poisson point process and have a duration that

is exponentially distributed. Interference with such properties has been found to be dominant in broadband PLC networks [104] and is expected to become significant in narrowband PLC networks with the dense deployment of smart meters based on multiple inoperable standards [73]. For these scenarios, I show that the Middleton class-A and Gaussian mixture models are the appropriate interference marginal distributions, a fact verified by prior empirical studies [26, 96].

In addition, I propose two empirical models to capture the temporal dependence in narrowband PLC networks and *bursty* interference in wireless transceivers. Using noise data collected in medium and low voltage sites, I propose a *cyclostationary* noise model that models the periodicity of the noise both in the temporal and spectral domain as an excitation of a filter bank by an AWGN signal. Similarly, using interference data collected from a laptop provided by Intel for the 2.4 GHz ISM band, I show that a hidden Markov model with Gaussian emissions (GHMM) provides a good fit for the inter-arrival times and durations of the bursty emissions exhibited by this data.

2. **Message-Passing OFDM Receivers in Impulsive Noise Channels:** This contribution builds on the results of the previous contribution which states that the effect of uncoordinated interference can be modeled as additive impulsive noise with a Gaussian mixture marginal and HMM temporal dynamics. Assuming that the statistics of the interference remain stationary for the duration of data transmission, the

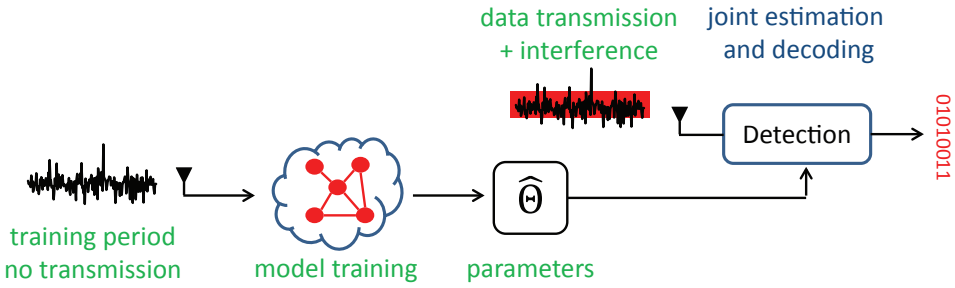


Figure 1.2: Training-based receivers: the parameters of the interference model are estimated during a quiet time and used in the receiver to perform detection. The interference statistics should be slowly varying otherwise the receiver might suffer from model mismatch.

interference model parameters are estimated from the noise samples during a *training* period, a quiet time when no data transmission occurs. The resulting interference models are then used to design receivers that mitigate the effect of this interference increasing the communication performance dramatically over traditional OFDM receivers. In particular, I propose two types of receivers: an Expectation-Maximization (EM) receiver that treats the presence of interference as a latent variable, and a message-passing receiver that adopts a Bayesian approach by treating the interference models as priors while attempting to jointly estimate the channel and the interference while decoding the transmitted data. The proposed receivers provide gains of up to 13dB over the typical DFT receiver at the same asymptotic complexity.

3. **Robust Message-Passing OFDM Receivers in Impulsive Noise Channels:** In this contribution, I consider two practical limitations

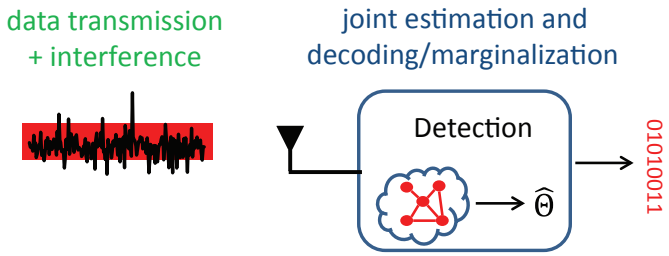


Figure 1.3: Robust receivers adapt their internal interference models based on the received data. As a result, they don't require extra training time and avoid the model mismatch problem in fast varying environments.

that the message-passing OFDM receivers proposed in the previous contribution might suffer from: (i) the lack of knowledge of the interference model parameters, (ii) and a mismatch in the assumed interference model. These limitations can occur if there is not enough quiet time for parameter training or if the receiver is operating in fast-varying environments where the noise statistics are changing rapidly. I propose two classes of *robust* receivers that are able to adapt to the interference environment. The first class does *blind* estimation of interference model parameters; i.e., it estimates the model parameters without an explicit training period. The second class uses the automatic relevance determination (ARD) prior, commonly used in sparse Bayesian learning and Bayesian compressed sensing, to design robust receivers that are independent of specific interference models leveraging only the noise sparsity in the time domain.

1.4 Organization

The dissertation is organized as follows:

Chapter 2 presents a brief overview of the basic concepts used in this work. It starts with the basic OFDM system model and how it is affected by interference that is non-AWGN. Then, it describes various techniques used to manage interference in communication networks highlighting some of their limitations that result in the presence of residual interference. Different statistical models are described that capture the characteristics of this interference in wireless networks. The chapter ends with a discussion of prior work on OFDM receiver design in uncoordinated interference and a brief overview of graphical models and their application to interference modeling.

Chapter 3 proposes various statistical models of uncoordinated interference in both PLC and wireless networks. It starts with statistical-physical modeling of uncoordinated interference in PLC networks using temporal Poisson point processes for emission arrivals. Then, it proposed two empirical models for temporal dependance that captures cyclostationarity in PLC noise and burstiness in wireless interference.

Chapters 4, 5, and 6 discuss how the models discussed in Chapter 3 can be utilized for OFDM receiver design. In particular, Chapter 4 describes an EM-based OFDM receiver that leverages the interference model to improve the communication performance under the constraint of independent channel decoding. Chapter 5 proposes a Bayesian inference framework for joint

channel/interference estimation and data decoding. The proposed receivers leverage recent advances in soft-input soft-output decoding, approximate message passing, and sparse signal recovery to design a low complexity reciever that is within 1dB from a lower bound. Chapter 6 builds on Chapter 5 to propose robust receivers that can be employed in rapidly varying interference environments.

Finally, Chapter 7 summarizes the contributions of this dissertation and outlines avenues for future research.

1.5 Notation

Vectors and matrices are denoted by boldface lower-case ($\mathbf{x}$) and upper-case notation ($\mathbf{X}$), respectively. $\mathbf{X}_{\mathcal{R},\mathcal{C}}$ then represents the sub-matrix constructed from rows $\mathcal{R}$ and columns $\mathcal{C}$ of $\mathbf{X}$, where the simplified notation $\mathbf{X}_{\mathcal{R}}$ means $\mathbf{X}_{\mathcal{R},:}$ and ":" indicates all columns of $\mathbf{X}$. The notations $(\cdot)^T$ and $(\cdot)^*$ denote transpose and conjugate transpose, respectively. The probability density function (pdf) of a random variable (RV) X is denoted by $p_X(x)$, with the subscript sometimes omitted when clear from the context. Similarly, for discrete RVs, the probability mass function (pmf) is denoted by $P_X(x)$. For a circular Gaussian RV with mean μ and variance γ, I write the pdf as $\mathcal{N}(x;\mu,\gamma)$. The expectation and variance of a RV are then given by $\mathbb{E}\{\cdot\}$ and $\mathbb{V}\{\cdot\}$, respectively. I use the *sans-serif* font to indicate frequency domain variables like X.

1.6 List of Acronyms

ADC: Analog to Digital Converter

AMP: Approximate Message Passing

AWGN: Additive White Gaussian Noise

ARD: Automatic Relevance Determination

ARMA: Auto Regressive Moving Average

BER: Bit Error Rate

BP: Belief Propagation

CGM: Cyclostationary Gaussian Model

CoMP: Coordinated Multi-Point

DFT: Discrete Fourier Transform

EM: Expectation Maximization

FFT: Fast Fourier Transform

ICI: Inter-Carrier Interference

ISM: Industrial, Scientific, and Medical band

IDFT: inverse Discrete Fourier Transform

JCNED: Joint Channel/Noise Estimation and Decoding

JNED: Joint Noise Estimation and Decoding

GAMP: Generalized Approximate Message Passing

GHMM: Gaussian Hidden Markov Model

GM: Gaussian Mixture

QAM: Quadrature Amplitude Modulation

LDPC: Low Density Parity Check

LMMSE: Linear Minimum Mean Square Estimate

LTE: Long Term Evolution

LTI: Linear Time Invariant

LPTV: Linear Periodically Time-Varying

LV: Low Voltage

MAC: Media Access Control

MAP: Maximum a-Posteriori Probability

MC: Markov Chain

MCA: Middleton Class-A

MIMO: Multiple Input Multiple Output

MMSE: Minimum Mean Squared Error

ML: Maximum Likelihood

MUD: Multi-user Detection

MSE: Mean Squared Error

MV: Medium Voltage

NBI: Narrowband Interface

NSI: Noise State Information

NMSE: Normalized Mean Squared Error

OFDM: Orthogonal Frequency Division Multiplexing

PEP: Pair-wise Error Probability

PGM: Probabilistic Graphical Model

PLC: Powerline Communications

PPP: Poisson Point Process

PRIME: PoweRline Intelligent Metering Evolution

SBL: Sparse Bayesian Learning

SC: Single Carrier

SER: Symbol Error Rate

SIC: Successive Interference Cancellation

SISO: Soft-Input Soft-Output

SNR: Signal to Noise Ratio

SP: Sum Product

STFT: Short time Fourier Transform

Chapter 2

Background

This chapter discusses the relevant background for multicarrier communication systems, interference management, graphical models and Bayesian inference using belief propagation. Section 2.1 provides the general system model for OFDM modulation and briefly discusses the impact of the impulsive noise on its performance. Section 2.2 briefly discusses interference management techniques that are typically employed in practice to reduce the interference level and highlight some of their shortcoming that are addressed by this dissertation. Section 2.3 discusses the statistical models used in the literature to model uncoordinated interference that cannot be mitigated by typical interference management techniques. Section 2.4 discusses prior work on designing OFDM receivers in impulsive interference and how they relate to receivers proposed by this dissertation. Finally, Section 2.5 ends with a brief overview of graphical models; their application to modeling OFDM systems and interference; and Bayesian inference using belief propagation.

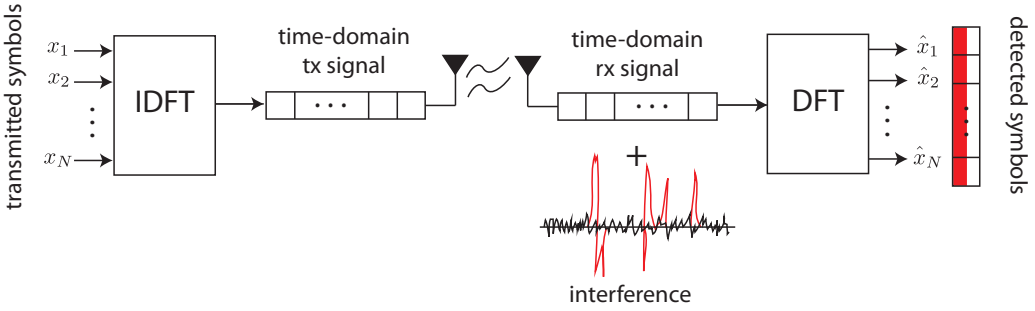


Figure 2.1: A simplified OFDM system model and its effect on impulsive interference: the DFT operation spreads the impulsive interference (in red) across all subcarriers.

2.1 OFDM Systems

Many modern wireless communications systems, such as IEEE 802.11n (Wi-Fi) and LTE (cellular), and recent powerline communication standards, such as PRIME and IEEE P1901.2, have adopted orthogonal frequency division multiplexing (OFDM) as their modulation scheme.

OFDM divides the transmission bandwidth into many subbands. The frequency at the center of each subband is called the subcarrier. Data is sent over many of the subbands at once. Pilot subcarriers carry known symbols that are used for channel estimation, while null subbands are loaded with zero power to adhere to out-of-band emissions requirements, relax cut-off filter requirements, or facilitate interference mitigation. Many OFDM standards also adapt the transmission in each subband based on the subband's signal-to-noise ratio (SNR) [37, 94] to increase the data throughput.

Most OFDM systems are based on the DFT and its implementation

using the fast Fourier transform (FFT) as illustrated in Figure 2.1. N data symbols are pre-coded using IDFT of length N. The resulting *time-domain* signal is transmitted through a frequency selective channel and corrupted by noise and interference present at the receiver. Then, the receiver takes the DFT of the received signal and applies minimum distance decoding on each symbol. These operations can be represented mathematically as

$$\mathbf{r} = \mathbf{H}\mathbf{F}^*\mathbf{x} + \mathbf{n} \xrightarrow{\text{DFT}} \mathbf{y} = \mathbf{F}\mathbf{r} = \mathsf{H} \circ \mathbf{x} + \mathbf{F}\mathbf{n}. \tag{2.1}$$

Here, $\mathbf{r}, \mathbf{x}$, $\mathbf{n}$, and H are $N \times 1$ complex vectors denoting the received OFDM signal, the transmitted symbols, the noise+interference vector, and frequency domain channel, respectively. $\mathbf{H}$ is a cyclic matrix formed by the vector $\mathbf{h}$ of the channel impulse response. By design, the OFDM receiver will diagonalize $\mathbf{H}$ into $\mathbf{\Lambda} = \mathbf{F}\mathbf{H}\mathbf{F}^* = \text{diag}\{\mathsf{H}\}$, where $\mathsf{H} = \mathbf{F}\mathbf{h}$ and $\circ$ denotes the Hadamard product. As a result, a sub-carrier k experiences a flat fading channel H_k; this significantly simplifies the equalization of the multipath channel. The described receiver is optimal under AWGN channels since the unitary transformation $\mathbf{F}$ does not change the statistics of the additive Gaussian noise. However, under non-AWGN interference, this receiver is highly suboptimal by tens of dBs [45–47].

To get a better understanding of the effect of impulsive interference on the communication performance, we consider the comparison of an OFDM system with a single carrier (SC) system given in Figure 2.2. The multiplication of the transmitted vector $\mathbf{x}$ by $\mathbf{F}^*$ spreads the "information" about each

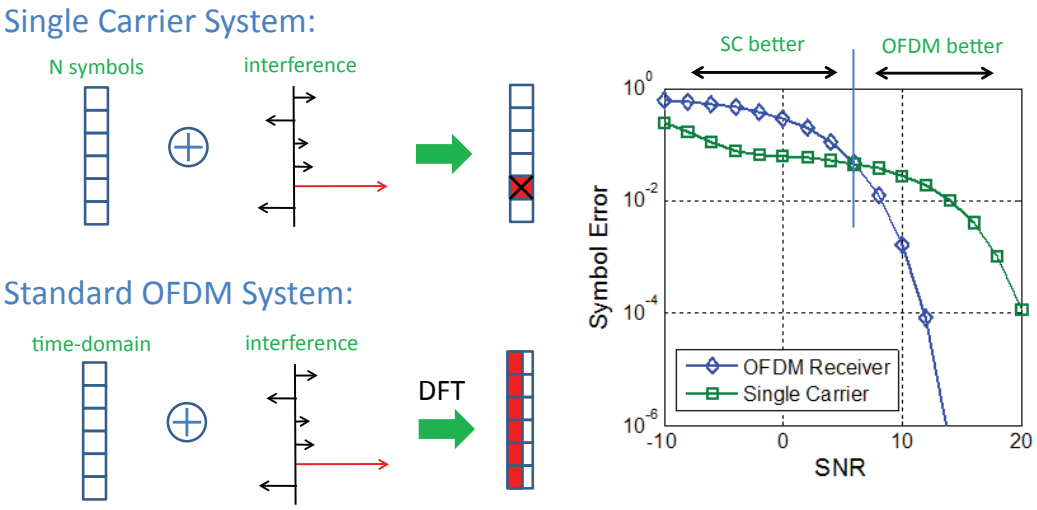


Figure 2.2: A simplified OFDM system model and its effect on impulsive interference.

symbol x_i across all of the OFDM symbol of duration N symbol times, providing a type of *code diversity* formalized for impulsive channels in [45, 47]. Alternatively, we can view the DFT operation at the receiver as smearing the impulse energy across all tones and reducing its impact at any given time sample. In contrast, in SC systems, the impulse energy will concentrate in time and severely affect the symbol being transmitted during at that instant. As a result, the symbol error rate curve in Figure 2.2 shows two regimes. In the first, the impulse energy is high enough to corrupt an entire OFDM symbol. However, since its effect in SC is localized, SC outperforms the DFT-based OFDM receiver. In the second, the impulse energy is low enough that when spread across the entire OFDM symbol its effect is significantly reduced and the DFT-based OFDM receiver outperforms the SC.

It should be noted the regime where SC outperforms OFDM is in a sense artificial: while the SC detector is MAP optimal, the DFT-based OFDM ignores the dependency of the noise and interference samples across sub-carriers. In fact, [78] shows that under a wide range of practical conditions, the mutual information between the decoded and the transmitted symbols is higher for SC than it is for a mutlicarrier system that performs independent detection across sub-carriers ignoring the dependencies. However, PEP-analysis performed in [45, 47] of the MAP OFDM decoder and the SC decoder shows that the MAP OFDM decoder outperforms the SC decoder by tens of dBs. The main drawback, discussed extensively in subsequent chapters, is that MAP decoding of OFDM in impulsive noise channels is exponential in computational complexity and, with number of tones ranging in the thousands for modern OFDM systems, intractable even for desktop simulations. *The main contribution of this dissertation is designing computationally efficient algorithms that approximate the MAP OFDM decoder recovering the performance loss suffered by typical OFDM receivers.*

2.2 Interference Management

A variety of interference management techniques has been developed to keep the interference power under control and ensure the scalability of communication networks. These techniques vary in the type of information they leverage for interference management. This information could encompass the deployed network topology or even the current state of the network and the

present interferers through coordination. Interference management techniques remain an active area of research both in industry and academia. A brief overview of such techniques is provided in the following.

2.2.1 Shielding

Shielding constitutes the most basic form of interference management techniques. By taking interference measurements across the platform, the sources of interference can be determined and shielded accordingly. This will protect the wireless transceiver from any emissions emitted from those sources. A drawback is that shielding is in general costly and heavy: something to avoid in designing modern ultra-light platforms. In addition, since shielding cannot be applied to other communicating devices or interferers outside the platform, it cannot protect against adjacent channel or out-of-platform interference[86].

2.2.2 Orthogonal Multiple Access Schemes

Such schemes allow multiple users to share a common propagation medium by making their transmissions orthogonal to each other in some dimension. Common dimensions include: time resulting in time division multiple access (TDMA), frequency resulting in frequency division multiple access (FDMA), or code resulting in code division multiple access (CDMA) [37, 94]. While TDMA and FDMA require coordination, CDMA can operate asynchronously [37, 94]. When perfect orthogonality is achieved, interference is canceled completely. However, in practice, orthogonality might not

be achieved due to timing offsets and channel effects as is the case in the CDMA uplink channel [94]. This results in residual interference. Furthermore, it was shown in [58] that there are fundamental limits on cooperation: an interference-limited network cannot be converted to a noise-limited one. Basically, due to the finite number of users that can orthogonalized, cooperation can only be achieved in limited size clusters. As a result significant, out-of-cluster interference is still present.

2.2.3 MAC Layer Access Schemes

Typically employed in decentralized networks like WiFi, MAC layer access schemes attempt to avoid simultaneous transmissions which would result in interference. Carrier sense multiple access (CSMA) is a popular access scheme that monitors the energy in the common medium to determine whether a transmission in underway. However, this procedure is prone to false decisions. As a result, this schemes cannot orthogonalize the access and inevitably result in residual interference.

2.2.4 Interference Cancellation

Interference cancellation techniques leverage the fact that interference seen from uncoordinated users is a communication signal. If this signal is strong enough, it can be successfully decoded at the receiver and subtracted from the received signal to improve its detection [37, 94]. Common methods of interference cancellation include multiuser detection (MUD), and successive

interference cancellation (SIC) [37, 94, 98]. Interference cancellation schemes are typically effective against dominant interferers whose power is high enough for successful decoding. However, signals from interferers whose power is below the successful decoding threshold or whose modulation is unknown will still appear as interference. Furthermore, this method is powerless against non-communication interferers such as many devices operating in ISM band and most interferers in powerline communication networks.

2.2.5 Precoding Techniques

MIMO precoding techniques leverage the freedom provided by multiple antennas to structure interference in a manner that reduces its impact on the different receivers. MIMO interference coordination has been an active area of research and many solutions has been proposed such as inter-cell interference cancellation [15], network MIMO (also known as CoMP) [35], and interference alignment [7]. The common limitation is the heavy reliance on coordination and sharing side information such as channel state information and user data (in the case of CoMP). This limits the applicability of such solutions in systems that lack a centralized or otherwise dedicated back-haul for sharing these information such as WiFi.

2.3 Statistical Modeling of Uncoordinated Interference

The previous section described how the presence of interference cannot be completely avoided in modern communication systems; however, as the

Model	Characterization	Application
Symmetric Alpha Stable	Characteristic function: $\Phi(\omega) = e^{-\sigma\|\omega\|^{\alpha}}$ α: characteristic exponent σ: dispersion	Sensor and *Ad hoc* networks
Gaussian Mixture	$p(x) = \sum_{k=0}^{M-1} \pi_k \mathcal{N}(x; 0, \gamma_k)$ M: number of components π_k: component probability γ_k: component variance	Cellular Networks with user clustering Two-tier Femtocell Networks
Middleton class-A	Gaussian Mixture with $M = \infty$, A: overlap index Γ: power ratio, $\pi_k = e^{-A}A^k/k!$ $\gamma_k = \gamma(k/A + \Gamma)/(1 + \Gamma)$ γ: variance	Dense WiFi Networks Cellular networks

Table 2.1: List of statistical-physical models for additive noise/interference and their applications in communication systems.

demand for higher data rates grows, it is becoming the performance bottle neck. As a result, statistical modeling of this uncoordinated interference is of great importance for receiver design and performance analysis. The two approaches for selecting an appropriate model are the *statistical-physical* and the *empirical* approaches. In the statistical-physical approach, one constructs a statistical model based on the physical principles governing the quantity of interest such as the spatial distribution of interferers, their power, and the channel fading statistics. The empirical approach uses experimental data to propose a statistical model that describes the measurements of interest.

2.3.1 Statistical-Physical Models for Uncoordinated Interference

Since the 1970s, significant effort has been devoted to developing statistical-physical models for impulsive noise based on the physical characteristics of the deployment environment. In his pioneering work, Middleton modeled these random spatio-temporal emissions, or the "noise field", using Poisson point processes (PPP) giving rise to the "Middleton class-A" and "Middleton class-B" noise models (for a recent review see [69]). More recently, this work has been extended to modeling field of interferers in many wireless and PLC networks using spatial and temporal PPPs. In [51, 87], it was shown that the interference from a homogeneous Poisson field of interferers distributed over the entire plane, an abstraction for wireless ad hoc networks, follows a symmetric alpha stable distribution. When we constraint the Poisson field of interferers to a finite region with or without guard zones, for example to model cellular systems, the generated interference follows a Middleton's class A distribution [41]. For Poisson clusters, in which each cluster contains Poisson distributed interferers, the interference follows a Gaussian mixture distribution [41]. This models cellular networks with user clustering and two-tier femtocell networks. These models are summarized in Table 2.1.

The temporal dependence of the interference in Poisson fields was investigated in [101] in which a model for second order statistics was proposed. More recently, a joint distribution of interference temporal samples in Poisson fields, with random interferer transmission duration, has been found to follow a multivariate Gaussian mixture distribution [40].

2.3.2 Empirical Models for Uncoordinated Interference

Empirical models are based on interference measurements collected in the field or the platform. For example, extensive measurement campaigns of terrestrial wireless installations indicate that the additive noise is impulsive, with peak noise amplitudes reaching up to 40 dB above the thermal background noise level [13, 56, 74, 83, 84]. The noise affecting powerline communications (PLC) has also been shown to be highly impulsive, as well as bursty [72, 73, 104]. These measurement campaigns were used to fit the interference data to different statistical models of the marginal pdfs such as Middleton's Class A [96], Nakagami-m [65], and Rayleigh [18] distributions empirically without considering the underlying physical models of interference generation. In the case when field data is not available, simulated data can be used instead. For example, recent work in [26] uses simulated data of filtered interference to support the Gaussian mixture and Middleton class-A models. Due to their experimental nature, these marginal empirical models and their estimated parameters might not generalize across different scenarios. As a result, these models are considered subpar to the statistical-physical models described in Section 2.3.1. That said, empirical models shine in modeling temporal dependencies as it is usually challenging to derive closed-form analytical joint distributions of interference's temporal samples from statistical-physical network models [40]. Examples of such tractable models include Hidden Markov models (HMM), Auto-Regressive (AR) models, and cyclostationary models that are commonly used to model PLC interference data [52, 104]. Table 2.2

Model	Characterization	Application
Gaussian Mixture	$p(x) = \sum_{k=0}^{M-1} \pi_k \mathcal{N}(x; 0, \gamma_k)$ M: number of components π_k: component probability γ_k: component variance	Marginal pdf of interference samples
Middleton class-A	Gaussian Mixture with $M = \infty$, A: overlap index Γ: power ratio, $\pi_k = e^{-A}A^k/k!$ $\gamma_k = \gamma(k/A + \Gamma)/(1 + \Gamma)$ γ: variance	Marginal pdf of interference samples
Rayleigh	$p(x) = \frac{x}{\sigma^2} e^{-x^2/\sigma^2}$ σ: mode	Marginal pdf of interference power
Nakagami-m	$p(x) = \frac{2m^m}{\Gamma(m)\Omega^m} x^{2m-1} e^{-mx^2/\Omega}$ m: shape Ω: spread	Marginal pdf of interference power
Gaussian Hidden Markov Model (GHMM)	$p(z_k \vert z_{k-1}, \cdots, z_0) = p(z_k \vert z_{k-1})$ $p(z_k = j \vert z_{k-1} = i) = \mathbf{T}_{ij}$ $p(x_j \vert z_j) = \mathcal{N}(x_j; 0, \gamma_{z_j})$ $\mathbf{T}$: state transition matrix γ_j: variance in state j	Temporal dependance
Autoregressive Moving-Average (ARMA)	$x_k = \epsilon_k + \sum_l \alpha_l x_{k-l} + \sum_j \theta_j \epsilon_{k-j}$ $p(\epsilon_k) = \mathcal{N}(\epsilon_k; 0, \gamma)$ $\{\alpha_l, \theta_j\}_{j,l}$: parameters γ: variance	Spectrally shaped interference (correlation between samples)
Cyclostationary	$p(x_i + T, \cdots, x_j + T) =$ $p(x_i, \cdots, x_j)$ T: period	Repeating Phenomena

Table 2.2: List of empirical models for additive noise/interference and their applications in communication systems.

summarizes the empirical models and their applications.

2.4 Prior Work on the Design of Interference-Limited Receivers

Motivated by huge potential gains, there has been significant amount of prior work that attempts to redesign the OFDM receiver taking into account the impulsive nature of uncoordinated interference. For simplicity of exposition, I categorize the prior work into three categories: time-domain preprocessing, sparse impulse noise reconstruction, and iterative receivers.

2.4.1 Time-Domain Preprocessing Techniques

These techniques use nonlinear estimators or sample thresholding to mitigate the effect of the impulse noise in the time-domain before passing it to the conventional DFT receiver. By treating the time-domain OFDM sample at time t (u_t), as being a Gaussian random variable, Haring derives $\mathbb{E}\{u_t|r_t\}$, its MMSE estimate given the received sample r_t and the impulsive noise model, and passes it to the DFT receiver for decoding [45]. Thresholding techniques, on the other hand, compare each OFDM time-domain sample to a threshold T: if a sample exceeds T, it is either blanked or clipped. The analysis and various threshold selection methods for such techniques are given in [95, 103]. While having a low implementation complexity, these techniques don't exploit the signal space diversity provided by the OFDM modulation [47]; thus, their performance deteriorates when the power of impulses is close

to the power of the OFDM signal or when higher order modulations are used [45]. Under these conditions, the receiver is more error prone to threshold an actual OFDM signal sample that is uncorrupted by an impulse. Furthermore, these techniques don't account for the variation in the received signal due to a fading dispersive channel which severely limits their application in practical systems.

2.4.2 Sparse Impulsive Noise Reconstruction

These methods assume that the impulsive noise is sparse in the time domain, and attempt to reconstruct it using the received signal on known tones (either null or pilot tones). The reconstructed impulse vector is then subtracted from the received signal and the result is passed to a DFT receiver. Exploiting the similarities between OFDM modulation with known tones and Reed-Solomon coding, Wolf proposed the use of frequency algebraic interpolation techniques to estimate the impulse vector [99]; this approach was later extended to more general settings in [2, 3]. Recently, recognizing the limitations of these methods in the presence of background noise, Caire proposed the use of compressed sensing for impulse noise reconstruction [17]. This method was later extended to bursty noise in [55]. For the typical number of known tones in OFDM systems, these techniques can reconstruct only very sparse impulse noise (one impulse in a 256-tone OFDM system with 30 known tones) and their performance degrades significantly under more practical impulsive noise scenarios [17, 57]. A more robust approach based on sparse Bayesian

learning (SBL) was given in [57]. The SBL receiver not only achieves significantly better performance under practical impulsive noise scenarios but also enables the receiver to use all the tones, including data tones, in a joint noise estimation and data detection scheme. A limitation of these techniques is that they still depend on linear channel estimation to utilize the pilot tones which is suboptimal in impulsive noise, and, with the exception of the SBL receiver, don't use the data tones or perform joint data detection and impulse noise estimation. However, the main limitation is that these methods don't utilize the statistical models for interference discussed in the previous section.

2.4.3 Iterative Receivers

These receivers alternate between the time and frequency domains to mitigate the effect of impulsive noise. In a given iteration, a time-domain pre-processing technique, such as clipping, can be applied to the received signal followed by symbol detection and time-domain correction for the following iteration [66, 102]. While low in implementation complexity, the limitations of the time-domain preprocessing methods carry over to these receivers. Another approach is to leverage the central limit theorem and the resulting approximate Gaussian behavior of the impulsive noise and the OFDM signal in the frequency and time domains, respectively, to design linear estimators and detectors in each domain and sequentially apply them reducing the effect of impulsive noise with each iteration [46, 71]. The main limitations of these techniques is that they don't incorporate channel estimation. In addition, although

[46] achieves significant performance gains, it requires setting parameters by simulation which limits its application in practical systems.

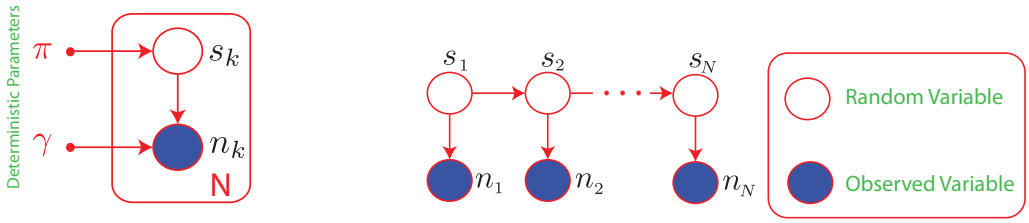
2.5 Probabilistic Graphical Models

Probabilistic graphical models (PGM) are probabilistic models for which a graph represents the conditional independence structure between the random variables or, equivalently, the factorization properties of the joint distribution [12]. This graphical representation enables easy understanding of the relationships between the random variables and, as seen later in this section, allows for efficient inference using message passing algorithms. This section introduces PGM basics required by later chapters and discusses how they can be applied to modeling uncoordinated interference in OFDM systems.

2.5.1 Graphical Representation of Interference Models

Section 2.3 discusses various models that have been proposed in the literature to model uncoordinated interference in communication systems. Due to their generality and tractability, we restrict our attention to Gaussian mixture models and hidden Markov models. In particular, Middleton class-A and symmetric alpha stable models can be approximated as a Gaussian mixture distribution [54, 97].

As discussed, a random variable $\mathbf{X}$ has a Gaussian mixture distribution if its probability density function (pdf) is a weighted sum of Gaussian



(a) Gaussian Mixture

(b) Hidden Markov

Figure 2.3: Graphical model representation of the interference distribution for two cases: a) i.i.d. Gaussian mixture pdf and b) a Hidden Markov Model with Gaussian emission density.

distributions $\mathcal{N}\left(\mu_k, \sigma_k^2\right), 1 \leq k \leq K$ given by

$$p_X\left(x\right) = \sum_{k=1}^{K} \pi_k \cdot \mathcal{N}\left(x; \mu_k, \sigma_k^2\right), \tag{2.2}$$

where $\mathcal{N}\left(x; \mu_k, \sigma_k^2\right)$ denotes the complex Gaussian distribution with mean μ_k and variance σ_k^2, and π_k is the mixing probability of the k-th Gaussian component. This distribution admits a simple latent variable model interpretation with the latent variable $\mathbf{S}$ indicating which Gaussian component was used to generate a particular realization of $\mathbf{X}$. The probability mass function (pmf) of $\mathbf{S}$ is given by the mixing vector $\boldsymbol{\pi} = [\pi_1 \cdots \pi_K]$. For interference in communication systems, the mean is typically assumed to be zero; otherwise, it can be subtracted from the received signal; i.e. we set $\mu_k = 0, \forall k$. A graphical representation of the Gaussian mixture model is given in Figure 2.3a. Since the latent variables in the Gaussian mixture model are independent, there is no temporal dependence between the noise samples. Hidden Markov models, on the other hand, introduce temporal dependence between the noise samples by

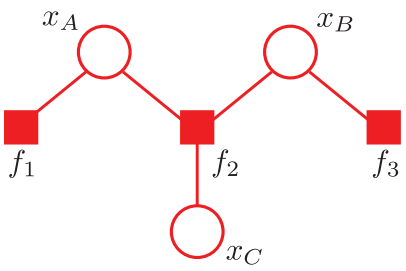


Figure 2.4: A factor graph representing the distribution in (2.3) with $f_1 = p(x_A)$, $f_2 = p(x_C|x_A, x_B)$, and $f_3 = p(x_B)$.

imposing a Markov structure on the hidden variables. The graphical representation of the resulting distribution is given in Figure 2.3b. In particular, each $\mathbf{S_i}$, that controls from which Gaussian component does the sample n_i come from, depends only on its previous value $\mathbf{S_{i-1}}$ and the transition probabilities of the governing Markov chain. The emission probabilities of the HMM model is Gaussian with variance determined by the latent variable $\mathbf{S_i}$.

2.5.2 Factor Graphs

Factor graphs are undirected graphical models that represent how a probability distribution factorizes. Each factor graph is composed of two types of nodes: variable nodes, typically represented by a circle, representing the random variables of the joint distribution; and factor nodes, represented by squares, denoting the individual factors in the factorization of the joint distribution. Edges exist strictly between variable and factor nodes to indicate what variables does each factor depend on. For example, the following factorization

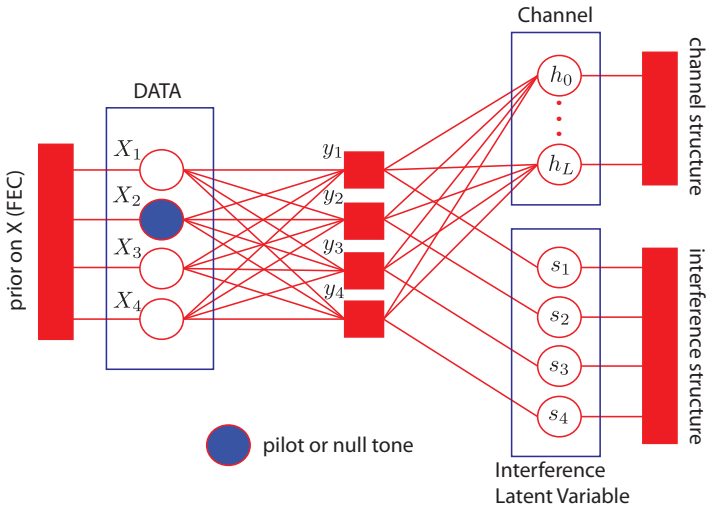


Figure 2.5: The graphical representation of the statistical dependencies of the variables in an OFDM system ($N = 4$).

of a given pdf

$$p\left(x_A, x_B, x_C\right) = p\left(x_A\right) p\left(x_B\right) p\left(x_C | x_A, x_B\right) \tag{2.3}$$

can be expressed by the factor graph given in Figure 2.4.

2.5.3 Factor Graph Representation of the OFDM System Model

Section 2.1 described the mathematical model behind OFDM systems. An alternate description of the system can be provided using a factor graph representation. The three main types of variables in this graph are: the transmitted data $\mathbf{x}$, the channel taps $\mathbf{h}$, and the interference latent variables $\mathbf{s}$. Figure 2.5 represents the probabilistic relationships between the variables in

an OFDM system by the following factorization of the joint probability density

$$p\left(\mathbf{x},\mathbf{s},\mathbf{h}|\mathbf{y}\right) \propto p\left(\mathbf{x}\right) p\left(\mathbf{h}\right) p\left(\mathbf{s}\right) \prod_{i=1}^{N} p\left(\mathbf{y_i}|\mathbf{x},\mathbf{s},\mathbf{h}\right). \tag{2.4}$$

Additional modeling assumptions on $\mathbf{x}, \mathbf{h}$, and $\mathbf{s}$ will lead to additional factorizations on their priors $p\left(\mathbf{x}\right), p\left(\mathbf{h}\right)$, and $p\left(\mathbf{s}\right)$. Some of the data symbols are pilot or null tones with known values typically used for channel estimation and spectral efficiency in wireless and powerline communication systems. These variables are considered to be observed nodes and they are set to their known values (blue nodes in Figure 2.5). The factor nodes represent the structural and statistical relationships between those variables. For example, the factor node connecting the data variables can represent the FEC code that was used to encode the bits that were translated into the transmitted symbols $\mathbf{x}$. The channel-structure factor node can represent the structured sparsity typically observed in wireless channels [85]. The factor node connected to the latent variables of the interference represents the temporal dependence and the prior probability of each Gaussian component.

Chapter 3

Statistical Modeling of Uncoordinated Interference

Statistical models of interference are important for evaluating and optimizing the performance of communication systems both in wireless and powerline systems. The effectiveness of these models depends on their ability to capture the statistics of the underlying random properties of the interference relevant to communication, such as burst durations and inter-arrival times of the interference emissions. As discussed in Section 2.3, the main modeling approaches for uncoordinated interference are statistical-physical and empirical models. While different statistical-physical models for uncoordinated interference in wireless networks have been proposed, no such models have been derived for powerline communication networks (PLC). Furthermore, modeling temporal dependencies remains a challenging area of interference modeling. This chapter tries to address these shortcomings by deriving statistical-physical models of interference in PLC networks and proposing various empirical models to capture the temporal dependencies in both PLC and wireless systems.

3.1 Statistical-Physical Modeling of PLC Interference

Powerline distribution networks are increasingly being employed to support smart grid communication infrastructure and in-home LAN connectivity. However, their primary function of power distribution results in a hostile environment for communication systems. In particular, asynchronous impulsive noise, with levels as high as 50 dB above thermal noise, causes significant degradation in communication performance. Much of the prior work uses limited empirical measurements to propose a statistical model for instantaneous statistics of asynchronous noise. In this section, I derive a canonical statistical-physical model of the instantaneous statistics of asynchronous noise based on the physical properties of the PLC network, and validate the distribution using simulated and measured PLC noise data. The results of this section can be used to analyze, simulate, and mitigate the effect of the asynchronous noise on PLC systems.

3.1.1 Introduction

Powerline networks are increasingly employed for communication purposes. These purposes vary from Internet connectivity inside the house to supporting smart grid applications such as automatic meter reading, device-specific billing and smart energy management. These powerline communication networks (PLC), initially designed for power transfer, result in a hostile environment for communication systems. Reflections and temporal variations in the PLC channel and correlated impulsive noise are the two main

impairments for reliable communication [50]. This dissertation focuses on noise/interference statistics, and refers readers interested in channel modeling to [8, 39, 50, 91, 105]. The non-Gaussian noise in PLC networks can be categorized into three main categories: generalized background noise, periodic impulsive noise, and asynchronous impulsive noise [104]. The first type has an exponentially decaying power spectral density superimposed with narrowband interference, while the second consists of broadband impulses occurring periodically. On the other hand, the asynchronous impulsive noise consists of random impulses of varying durations. It is mainly caused by switching transients of various appliances and devices in individual homes and businesses present on the network [50, 104]. Additional interference can also be picked up by the PLC network acting as an antenna for wireless in-band and aliased signals [50]. This impulsive noise, with levels as much as 50 dB above thermal noise, is considered the main cause of errors in PLC communications [50].

Frequency domain empirical studies fit the spectrally shaped background noise to various spectral models [22, 49]. Likewise, time domain properties of the asynchronous impulsive noise, such as impulse inter-arrival times, impulse durations, and instantaneous statistics, have been experimentally investigated in [18, 22, 91, 104]. I refer the reader to [91] and [104] for modeling the impulse inter-arrival times and impulse durations.

This section focuses on the instantaneous amplitude statistics of the asynchronous impulsive noise which are important properties for communication system performance and simulation [6, 65, 96]. Prior work fits the noise

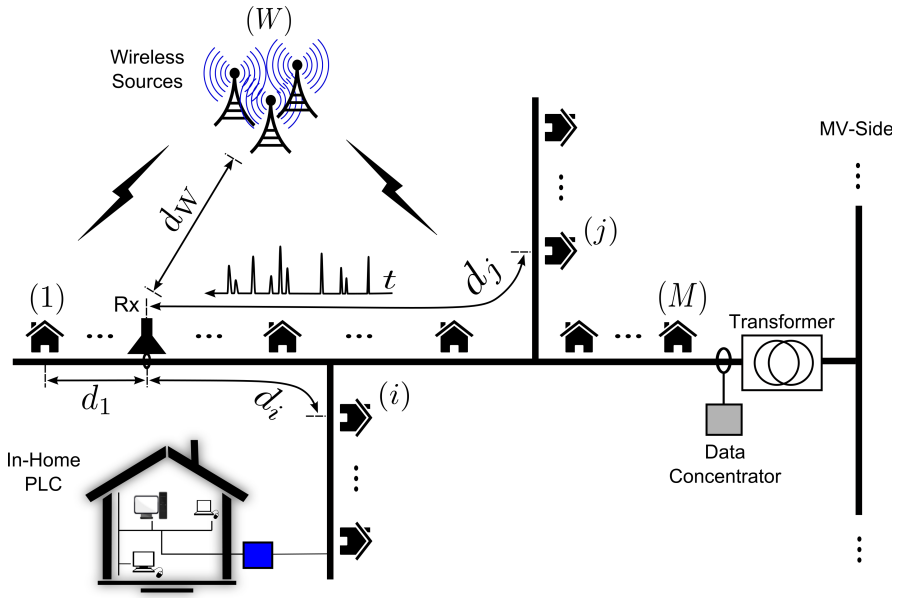


Figure 3.1: A system model for a low-voltage powerline communications network and an in-home PLC LAN with interference sources. Each interference source can be either on the powerline or from an external wireless source. Each interferer emits asynchronous impulsive noise at a distance d_m from the receiver for $m = 1, \ldots, M$.

data to different statistical models such as Middleton's Class A [96], Nakagami-m [65], and Rayleigh [18] distributions empirically without considering the underlying physical models of interference generation. Recent work in [26] supports the Gaussian mixture and Middleton noise models by filtering the interference through a PLC channel in a Monte Carlo simulation and studying the resulting statistics of the simulated noise.

3.1.2 Contribution

As a follow-up to previous work, I derive an analytical statistical-physical model of the first-order distribution of the asynchronous impulsive noise in PLC networks based on physical models of the PLC channel and the generated interference. Temporal and higher order statistics are left for future work. On top of that, I validate our models using Monte-Carlo simulations and experimental data collected on a PLC network.

Throughout this section, I use a slightly different notation that the rest of this dissertation to make the derivation easier to follow. In particular, random variables are represented using boldface notation, deterministic parameters are represented using non-boldface type, $\mathbb{E}\{f(\mathbf{X})\}$ denotes the expectation of the function $f(\mathbf{X})$ with respect to the random variable $\mathbf{X}$, and $\mathbb{P}(\cdot)$ denotes the probability of a random event.

3.1.3 System Model

I consider a power-distribution or an indoor PLC network in which a randomly located receiver receives a signal of interest in the presence of interfering signals. A typical system model for a low-voltage PLC network is given in Figure 3.1. In this model, there are M interferers that are a combination of various homes connected to a transformer and some wireless sources such as AM transmissions. The PLC environment is very dynamic and can exhibit different characteristics on hourly basis, such as variations in load impedance during the day period [50, 90, 104]. However, this section focuses on deriving

instantaneous impulse statistics on the time scale observed by a communication system. As a result, I ignore the large scale variations in the environment, and assume it to be stationary on the desired time scale. The interference experienced by a receiver at a reference time $t = 0$ due to emissions that arrived within a time interval of duration T from the reference time is given by

$$\mathbf{I}(T) = \sum_{m=1}^{M} \mathbf{I}_m(T) \tag{3.1}$$

where $\mathbf{I}_m(T)$ is the interference resulting from interference source m. Consequently, the interference due to all emissions that arrived in the past until time t is given by

$$\mathbf{\Psi} = \lim_{T \to \infty} \mathbf{I}(T). \tag{3.2}$$

Although taking $T \to \infty$ might seem to contradict the stationarity assumption mentioned earlier, I will show that in practice, due to an upper bound on the maximum interference duration, this is not the case and the result holds for the desired time scale. The objective is to find the first-order statistics of this total interference by calculating the characteristic function of $\mathbf{I}_m(T)$ for each interferer m. Toward this end, I focus on finding appropriate statistical models for the impulsive emissions based on interferer's temporal profiles based on experimental studies found in the literature (typically up to 20MHz).

3.1.3.1 Interference Emissions Modeling

Figure 3.2 shows the superposition of impulsive emissions due to source m. Each impulse i is made up of two parameters: an arrival time relative to

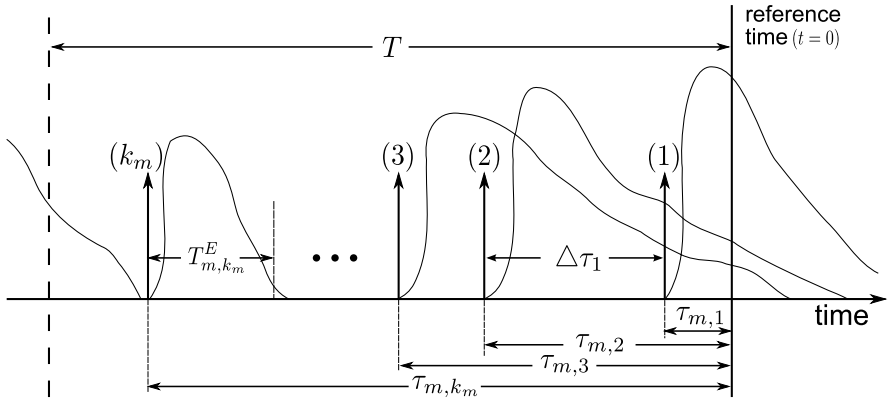


Figure 3.2: Superposition of impulses generated by source m: vertical arrows are illustrations indicating arrivals, k_m is the number of arrivals within time duration T, and $t = 0$ is the reference time.

the reference time denoted by $\boldsymbol{\tau}_{m,i}$ (indicated by an arrow) and an impulse duration denoted by $\mathbf{T}^E_{m,i}$. The dynamics of the emissions are captured by the inter-arrival times between impulses denoted by $\{\triangle\boldsymbol{\tau}_i : i \in \mathbb{N}\}$. Various experimental studies investigated the temporal properties of asynchronous impulsive emissions in PLC networks [18, 22, 91, 104]. In particular, measurements done in [91] and [104] showed that the inter-arrival time between two consecutive impulses fits an exponential distribution; i.e., the inter-arrival time between impulse i and impulse $i + 1$ has the following distribution

$$\triangle\boldsymbol{\tau}_{m,i} \sim \mathrm{Exp}\,(\lambda_m)$$

where λ_m is the emission *rate* of source m. Since two impulses arriving at the same time are indistinguishable (they add up constructively), the process $\boldsymbol{\Lambda}_m = \{\boldsymbol{\tau}_{m,i} : i \in \mathbb{N}\}$ representing the impulse arrival times for source m is a counting process with jumps of size one. This combined with the exponential

inter-arrival times, makes it a time Poisson point process with *rate* λ_m. As a result, the interference emissions in our model are characterized by a set of time Poisson point processes $\{\mathbf{\Lambda}_i(\lambda_i) : 1 \leq i \leq M\}$ corresponding to each interferer in Figure 3.1. This modeling can be generalized to indoor PLC networks where the interference sources are individual appliances [91]. On the other hand, the statistics of the impulse duration $\mathbf{T}^E_{m,i}$ have been studied in [22, 104]. It was found that a typical impulse has a duration ranging from about 10μs to $1m$s with a distribution that is loosely exponential and a typical value of hundreds of μs [104]. The exact distribution of the impulse duration is not important since the derivation depends only on its first moment $\mathbb{E}\left\{\mathbf{T}^E_{m,i}\right\}$.

3.1.3.2 Interference Channel Modeling

The PLC channel properties have been studied extensively in [8, 39, 105]. In [39], the PLC channel was fitted to a time-domain pulse model. On the other hand, [8] and [105] exploit the physical properties of the transmission line. The two-port network model presented in [8], represents each component of the PLC network, such as a cable or a transformer, by its equivalent two-port network description (ABCD or S-parameters). Then, transmission line (TL) theory is used to compute the equivalent channels and reflection impedances. On the other hand, the echo model, presented in [105], is simplified representation of the channel frequency response inspired by TL theory. The echo model describes the channel by the following equation

$$\mathbf{H}(f) = \sum_{j=1}^{N} \mathbf{g}_j e^{-\alpha(f)d_j} e^{-j2\pi f d_j/\nu} \tag{3.3}$$

where N is the number of paths, $\mathbf{g}_j$ is a random variable representing the reflection coefficient of each propagation path that depends on the observed load impedance, $\alpha(f)$ is the attenuation constant of the cables used, d_j is the length of each reflection path needs to travel, and ν is signal propagation speed through the wire. The impulse response of this channel has a delay spread τ_h between 1μs to around 4μs [39, 105].

The effect of this channel on an impulse i due to source m can be inferred by comparing the typical values of $\mathbf{T}_{m,i}^{E}$, the impulse durations presented in Section 3.1.3.1, to the channel delay spread τ_h given above. Since $\mathbf{T}_{m,i}^{E} \gg \tau_h$, the response of channel to the impulsive emission will have only one resolvable component and the channel will be a flat fading channel (the signal's propagation delay is much smaller than the impulse duration). A similar conclusion can be reached by looking at the channel's frequency response given in [39] and [105]. For an impulsive emission bandwidth between 1kHz and 100kHz corresponding to the aforementioned $\mathbf{T}_{m,i}^{E}$, the channel response is relatively flat. As a result, the discrete baseband equivalent channel of (3.3) is given by

$$\mathbf{h}[n] = \mathbf{h}e^{j\theta}e^{-\alpha_0 d}\delta[n] \tag{3.4}$$

where $\mathbf{h}$ is a random amplitude, $\boldsymbol{\theta}$ is a random phase uniformly distributed on $[0, 2\pi]$ under the uncorrelated fading assumption, $e^{-\alpha_0 d}$ is the path attenuation, d is the distance between the interferer and the receiver, and $\alpha_0 = \alpha(f_0)$ for some f_0 in the frequency band being considered (flat fading). Even if the channel exhibits frequency selectivity, the resulting multipath of the interfer-

ence can be lumped together into one longer impulse with a different amplitude distribution. The derivation depends only on the second order moment of the channel amplitude and thus can be applied to any channel distributions. For a wireless interferer, I assume a Rayleigh flat fading channel with pathloss proportional to $d^{-\gamma/2}$, where d is the distance of the source and γ is the pathloss exponent [94].

3.1.4 Statistical Modeling of $\mathbf{I}_m(T)$

Figure 3.2 shows a typical realization of impulse emissions within a window of duration T resulting from interference source m. The resulting interference at the receiver at a reference time $t = 0$, $\mathbf{I}_m(T)$, can be represented as

$$\mathbf{I}_m(T) = \gamma(d_m) \sum_{i=1}^{\mathbf{k}_m} \mathbf{h}_{m,i} e^{j\boldsymbol{\theta}_{m,i}} \mathbf{X}_{m,i} \tag{3.5}$$

where $\mathbf{k}_m$ is the number of impulses that arrived within a window of duration T, $\mathbf{h}_{m,i}e^{j\boldsymbol{\theta}_{m,i}}$ is the flat channel gain (based on (3.4)) between the interference source m and the receiver as seen by impulse i, and $\gamma(d_m)$ is the path attenuation. From Section 3.1.3.2, the channel attenuation can be expressed as

$$\gamma(d_m) = \begin{cases} d_m^{-\eta/2} & \text{if } m \text{ is a wireless source} \\ e^{-\alpha_0 d_m} & \text{if } m \text{ is a wired source} \end{cases} \tag{3.6}$$

where d_m is the distance between the source and the receiver. On the other hand, $\mathbf{X}_{m,i}$ is the random emission due to the duration of impulse i and can be represented as

$$\mathbf{X}_{m,i} = \mathbf{B}_{m,i} e^{j\boldsymbol{\phi}_{m,i}} \mathbf{1}\left(\boldsymbol{\tau}_{m,i} \leq \mathbf{T}^E_{m,i}\right) \tag{3.7}$$

where $\mathbf{1}(\cdot)$ is the indicator function, and $\mathbf{B}_i e^{j\boldsymbol{\phi}_i}$ represents the result of narrowband filtering of interference emissions performed at the receiver. The condition inside of the indicator function guarantees that the emission corresponding to impulse i is still active at the reference time $t = 0$ (See Figure 3.2). For example, in Figure 3.2 impulse k_m no longer has an effect at $t = 0$ while impulse 1 is still active as reflected in the indicator function's condition. $\mathbf{B}_i$ is an *i.i.d.* envelope and $\boldsymbol{\phi}_i$ is a random phase uniformly distributed on $[0, 2\pi]$. This representation is valid as long as $\mathbf{T}^E_{m,i} \gg \frac{1}{\triangle f_R}$ where $\triangle f_R$ is the receiver bandwidth [67]. This is the case for the values of $\mathbf{T}^E_{m,i}$ mentioned in Section 3.1.3.1, especially for broadband PLC ($\triangle f_R \approx 1$MHz). Expanding $\mathbf{I}_m(T)$ into its complex form, I obtain

$$\mathbf{I}_m(T) = \sum_{i=1}^{\mathbf{k}_m} \mathbf{h}_{m,i}\mathbf{B}_{m,i}\mathbf{1}\left(\boldsymbol{\tau}_{m,i} \leq \mathbf{T}^E_{m,i}\right) \times \left[\cos\left(\boldsymbol{\phi}_{m,i} + \boldsymbol{\theta}_{m,i}\right) + j\sin\left(\boldsymbol{\phi}_{m,i} + \boldsymbol{\theta}_{m,i}\right)\right]. \quad (3.8)$$

From (3.8), the joint characteristic function of the in-phase and quadrature-phase components of $\mathbf{I}_m(T) = \mathbf{I}^{(I)}_m(T) + j\mathbf{I}^{(Q)}_m(T)$, with implicit dependence on T, is given by

$$\Phi_{\mathbf{I}_m}(\omega) = \mathbb{E}_{\mathbf{I}_m}\left\{e^{j\omega_I \mathbf{I}^{(I)}_m + j\omega_Q \mathbf{I}^{(Q)}_m}\right\}$$
$$= \mathbb{E}\left\{e^{j\sum\limits_{i=1}^{\mathbf{k}_m} \mathbf{h}_{m,i}\mathbf{B}_{m,i}\mathbf{1}\left(\boldsymbol{\tau}_{m,i}\leq \mathbf{T}^E_{m,i}\right)|\omega|\cos\left(\boldsymbol{\phi}_{m,i}+\boldsymbol{\theta}_{m,i}+\omega_\phi\right)}\right\}$$

where $\mathbf{I}_m = \left[\mathbf{I}^{(I)}_m, \mathbf{I}^{(Q)}_m\right]^T$ and $\omega = [\omega_I, \omega_Q]^T$, $|\omega| = \sqrt{\omega_I^2 + \omega_Q^2}$, and $\omega_\phi = \tan^{-1}\left(\frac{\omega_Q}{\omega_I}\right)$. The expectation in the above equation is with respect to $\mathbf{k}_m$,

$\left\{\mathbf{B}_{m,i}, \mathbf{h}_{m,i}, \boldsymbol{\tau}_{m,i}, \boldsymbol{\phi}_{m,i}, \boldsymbol{\theta}_{m,i}, \mathbf{T}^{E}_{m,i} : 1 \leq i \leq \mathbf{k}_m\right\}$. Taking the expectation over $\mathbf{k}_m$, I obtain

$$\Phi_{\mathbf{I}_m}(\omega) = \sum_{k_m=0}^{\infty} \mathbb{P}\left(k_m \text{ arrivals in duration } T\right) \times \mathbb{E}\left\{ e^{j \sum_{i=1}^{k_m} \mathbf{h}_{m,i}\mathbf{B}_{m,i}\mathbf{1}\left(\boldsymbol{\tau}_{m,i} \leq \mathbf{T}^{E}_{m,i}\right)|\omega| \cos\left(\boldsymbol{\phi}_{m,i}+\boldsymbol{\theta}_{m,i}+\omega_{\phi}\right)} \mid k_m \right\} \quad (3.9)$$

Since $\boldsymbol{\Lambda}_m(\lambda_m)$ is a homogeneous Poisson time-point process, the number of impulse arrivals $\mathbf{k}_m$ in the window of duration T is Poisson distributed with distribution

$$\mathbf{k}_m \sim \text{Pois}\left(\lambda_m T\right).$$

Furthermore, given $\mathbf{k}_m$, the impulse arrival times $\{\boldsymbol{\tau}_{m,i} : 1 \leq i \leq \mathbf{k}_m\}$ are mutually independent and uniformly distributed on $[0, T]$; thus

$$\boldsymbol{\tau}_{m,i} \mid \mathbf{k}_m \sim \mathcal{U}(0, T) \text{ for } 1 \leq i \leq \mathbf{k}_m. \quad (3.10)$$

Assuming $\left\{\mathbf{B}_{m,i}, \mathbf{h}_{m,i}, \boldsymbol{\phi}_{m,i}, \boldsymbol{\theta}_{m,i}, \mathbf{T}^{E}_{m,i} \mid \mathbf{k}_m : 1 \leq i \leq k_m\right\}$ are all *i.i.d.* (i.e. statistically identical emissions for each impulse i), I can drop the index i and write (3.9) as

$$\begin{aligned}\Phi_{\mathbf{I}_m}(\omega) = \sum_{k_m=0}^{\infty} & \frac{e^{-\lambda_m T}\left(\lambda_m T\right)^{k_m}}{k_m!} \\ & \times \left(\mathbb{E}\left\{e^{j|\omega|\mathbf{h}_m\mathbf{B}_m\mathbf{1}\left(\boldsymbol{\tau}_m \leq \mathbf{T}^{E}_{m}\right)\cos\left(\boldsymbol{\phi}_m+\boldsymbol{\theta}_m+\omega_{\phi}\right)}\right\}\right)^{k_m} \\ & = e^{\lambda_m T\left(\mathbb{E}\left\{e^{j|\omega|\mathbf{h}_m\mathbf{B}_m\mathbf{1}\left(\boldsymbol{\tau}_m \leq \mathbf{T}^{E}_{m}\right)\cos\left(\boldsymbol{\phi}_m+\boldsymbol{\theta}_m+\omega_{\phi}\right)}\right\}-1\right)} \end{aligned} \quad (3.11)$$

Denoting the expectation in (3.11) by $\psi_{\mathbf{I}_m}(\omega)$, I obtain

$$\begin{aligned}\psi_{\mathbf{I}_m}(\omega) &\triangleq \mathbb{E}\left\{e^{j|\omega|\mathbf{h}_m\mathbf{B}_m\mathbf{1}\left(\boldsymbol{\tau}_m\leq\mathbf{T}_m^E\right)\cos\left(\boldsymbol{\phi}_m+\boldsymbol{\theta}_m+\omega_\phi\right)}\right\}\\ &\overset{(a)}{=} \mathbb{E}\left\{\left(1-\frac{\mathbf{T}_m^E}{T}\right)e^0+\frac{\mathbf{T}_m^E}{T}e^{j|\omega|\mathbf{h}_m\mathbf{B}_m\cos\left(\boldsymbol{\phi}_m+\boldsymbol{\theta}_m+\omega_\phi\right)}\right\}\\ &\overset{(b)}{=} 1-\frac{\mu_m}{T}+\frac{\mu_m}{T}\mathbb{E}\left\{e^{j|\omega|\mathbf{h}_m\mathbf{B}_m\cos\left(\boldsymbol{\phi}_m+\boldsymbol{\theta}_m+\omega_\phi\right)}\right\}\end{aligned} \tag{3.12}$$

where step (a) follows from taking the expectation over $\boldsymbol{\tau}_m$, and step (b) from taking the expectation over $\mathbf{T}_m^E$ with the notation $\mu_m = \mathbb{E}\left\{\mathbf{T}_m^E\right\}$. In step (a), I made the implicit assumption that $T > \mathbf{T}_m^E(\omega), \forall\omega \in \Omega$ where Ω is the probability space. This assumption is valid since in practice $\mathbf{T}_m^E$, the impulse duration, is bounded and follows a truncated distribution [104]. Further, (3.2) shows that I are interested in the limit as $T \to \infty$ which justifies our assumption. By using the identity

$$e^{ja\cos(\theta)} = \sum_{k=0}^{\infty} j^k \epsilon_k J_k(a)\cos(k\theta) \tag{3.13}$$

where J_k is the Bessel function of the k-th order, $\epsilon_0 = 1$ and $\epsilon_k = 2$ for $k \geq 1$, (3.12) can be written as

$$\begin{aligned}\psi_{\mathbf{I}_m}(\omega) = 1-\frac{\mu_m}{T}+\frac{\mu_m}{T}\mathbb{E}\Bigg\{&\sum_{k=0}^{\infty} j^k\epsilon_k J_k\left(|\omega|\,\mathbf{h}_m\mathbf{B}_m\right)\\ &\times\cos\left(k\left(\boldsymbol{\phi}_m+\boldsymbol{\theta}_m+\omega_\phi\right)\right)\Bigg\}.\end{aligned} \tag{3.14}$$

Since $\boldsymbol{\phi}_m$ and $\boldsymbol{\theta}_m$ are uniformly distributed on $[0, 2\pi]$, $\mathbb{E}_{\phi_m,\theta_m}\left\{\cos\left(k\left(\boldsymbol{\phi}_m+\boldsymbol{\theta}_m+\omega_\phi\right)\right)\right\} = 0$ for $k \geq 1$, and (3.14) reduces to

$$\psi_{\mathbf{I}_m}(\omega) = 1-\frac{\mu_m}{T}+\frac{\mu_m}{T}\mathbb{E}_{\mathbf{h}_m,\mathbf{B}_m}\left\{J_0\left(|\omega|\,\mathbf{h}_m\mathbf{B}_m\right)\right\}. \tag{3.15}$$

An approximation for the expectation term in (3.15) is given by

$$\mathbb{E}_{\mathbf{h}_m,\mathbf{B}_m}\left\{J_0\left(|\omega|\,\mathbf{h}_m\mathbf{B}_m\right)\right\} = e^{\frac{-|\omega|^2\mathbb{E}\left\{\mathbf{h}_m^2\mathbf{B}_m^2\right\}}{4}}\left(1+\Theta\left(|\omega|^4\right)\right) \tag{3.16}$$

where $\Theta\left(\|\omega\|^4\right)$ denotes a correction term with the lowest power of $\|\omega\|$ being four [67]. Fourier analysis shows that the behavior of the characteristic function in the neighborhood of zero governs the tail probabilities of the random variable. As a result, $\Theta\left(\|w\|^4\right) \ll 1$ for $\|w\| \to 0$, and can be ignored from (3.16) for modeling tail probabilities. For $\mathbf{h}_m\mathbf{B}_m$ Rayleigh distributed, $\Theta\left(\|w\|^4\right) = 0$ and the following result is exact. Substituting (3.16) into (3.15) and then into (3.11), I obtain

$$\begin{aligned}\Phi_{\mathbf{I}_m}(\omega) &= e^{\lambda_m\mu_m\left(-1+e^{\frac{-\|w\|^2\mathbb{E}\left\{\mathbf{h}_m^2\mathbf{B}_m^2\right\}}{4}}\right)} \\ &= e^{-\lambda_m\mu_m}\sum_{k=0}^{\infty}\frac{\left(\lambda_m\mu_m\right)^k}{k!}e^{\frac{-k\|w\|^2\mathbb{E}\left\{\mathbf{h}_m^2\mathbf{B}_m^2\right\}}{4}}\end{aligned} \tag{3.17}$$

where the second step follows from the Taylor expansion of the exponential function. Two important observations can be made about (3.17): 1) there is no dependence on T, and (2) it is the characteristic function of a Middleton Class A distribution with parameters given by

$$A_m = \lambda_m\mu_m = \lambda_m\mathbb{E}\left\{\mathbf{T}_m^E\right\} \tag{3.18}$$

$$\Omega_m = \frac{A_m \times \mathbb{E}\left\{\mathbf{h}_m^2\mathbf{B}_m^2\right\}}{2} = \frac{A_m\gamma\left(d_m\right)\mathbb{E}\left\{\mathbf{g}_m^2\mathbf{B}_m^2\right\}}{2} \tag{3.19}$$

where A_m is the overlap index that indicates the amount of impulsiveness of the interference originating from source m, and Ω_m is its mean intensity. The independence of (3.17) from T is important for deriving the statistics of the total interference and is discussed in the following section.

3.1.5 Statistical Modeling of the Total Interference $\mathbf{\Psi}$

The total interference as seen by the receiver is the superposition of impulses resulting from all available interference sources. Further, it should encompass the contribution of the impulse durations potentially spanning infinitely in the past. This is reflected in (3.2). However, as seen from (3.17), the statistics of the total interference $\mathbf{\Psi}$ depend only on the impulses that arrived within the maximum impulse duration which is finite. This has a simple intuitive explanation: any impulse that arrived before the maximum impulse duration would have died out by the reference time $t = 0$. On top of that, the maximum impulse duration is on the order of milliseconds (only 1% of total impulses exhibit a duration exceeding 1ms [104]). This duration is much lower than the rate of variation in the PLC environment which is on the order of hours and days [90]. This justifies the stationarity assumption mentioned in Section 3.1.3.

Let $\xi = \max\left\{\mathbf{T}_m^E(\omega) : \omega \in \Omega, 1 \leq m \leq M\right\}$, then (3.2) can be expressed as

$$\mathbf{\Psi} = \lim_{T\to\infty} \mathbf{I}(T) = \mathbf{I}(\xi) = \sum_{m=1}^{M} \mathbf{I}_m(\xi). \tag{3.20}$$

Assuming that impulses from different interference sources are independent and using the result from (3.17), I can express the characteristic function of the total interference as

$$\Phi_\Psi(\omega) = e^{-\sum_{m=1}^{M}\lambda_m\mu_m^E} \sum_{k_1=0}^{\infty}\cdots\sum_{k_M=0}^{\infty}\prod_{m=0}^{M}\frac{\left(\mu_m^E\lambda_m\right)^{k_m}}{k_m!} \times e^{-\|\omega\|^2\sum_{m=1}^{M}k_m\gamma(d_m)\mathbb{E}\left\{\mathbf{g}_m^2\mathbf{B}_m^2\right\}/4}. \tag{3.21}$$

This is the characteristic function of a Gaussian mixture distribution. Truncating the each infinite summation into N terms, (3.21) can be simplified into the more familiar form

$$\Phi_{\Psi}(\omega) = \sum_{i=1}^{N^M} \pi_i e^{-\|\omega\|^2 \sigma_i^2} \tag{3.22}$$

where

$$\pi = \begin{bmatrix} \frac{\lambda_1^0 e^{-\mu_1 \lambda_1}}{0!} \\ \vdots \\ \frac{\lambda_1^N e^{-\mu_1 \lambda_1}}{N!} \end{bmatrix} \otimes \cdots \otimes \begin{bmatrix} \frac{\lambda_M^0 e^{-\mu_M \lambda_M}}{0!} \\ \vdots \\ \frac{\lambda_M^N e^{-\mu_M \lambda_M}}{N!} \end{bmatrix}$$

and

$$\sigma^2 = \frac{1}{4} \begin{bmatrix} 0 \cdot \mathbb{E}\{\mathbf{h}_1^2 \mathbf{B}_1^2\} \\ \vdots \\ N \cdot \mathbb{E}\{\mathbf{h}_1^2 \mathbf{B}_1^2\} \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 \cdot \mathbb{E}\{\mathbf{h}_M^2 \mathbf{B}_M^2\} \\ \vdots \\ N \cdot \mathbb{E}\{\mathbf{h}_M^2 \mathbf{B}_M^2\} \end{bmatrix}$$

where $\pi = \begin{bmatrix} \pi_1 & \cdots & \pi_{N^M} \end{bmatrix}$ and $\sigma^2 = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{N^M}^2 \end{bmatrix}$. The operations $\otimes$ and $\oplus$ denote the Kronecker multiplication and sum respectively. These equations can be made arbitrary accurate by increasing N; however, 2 to 3 terms are usually sufficient in practice [61]. The amplitude distribution of the total interference can be deduced from (3.22) and written as

$$f_{\|\Psi\|}(\zeta) = \sum_{i=1}^{N^M} \pi_i e^{-\|\omega\|^2 \sigma_i^2} \frac{\zeta}{\sigma_i^2} e^{-\zeta^2/\sigma_i^2}$$

which is a sum of Rayleigh distributions.

3.1.6 Discussion

Eq. (3.21) and (3.22) describe the interference statistics under the general conditions given in Figure 3.1. These equations can be further simplified

Table 3.1: Statistical-physical modeling of asynchronous impulsive noise in different PLC networks. For each interfering source m, λ_m is the emission rate, μ_m is the mean, and d_m is the distance to the receiver. There are M interfering sources.

Scenario	**Model**
General PLC network $\{\lambda_m, \mu_m, d_m : 1 \leq m \leq M\}$	Gaussian Mixture π, σ^2 in (3.22)
One Dominant Interference Source λ, μ, d	Middleton's Class A $A = \lambda\mu, \Omega = \frac{A\gamma(d)\mathbb{E}\left\{\mathbf{h}^2\mathbf{B}^2\right\}}{2}$
Homogeneous PLC network $\lambda_m = \lambda, \mu_m = \mu, \gamma(d_m) = \gamma$ $\forall m \in \{1, \cdots, M\}$	Middleton's Class A $A = M\lambda\mu, \Omega = \frac{\lambda\mu\gamma\mathbb{E}\left\{\mathbf{h}^2\mathbf{B}^2\right\}}{2}$

by assuming more homogeneous environments with similar properties such as emission rates and channel and emission's amplitudes statistics. For example, environments with one dominant interference source will follow a Middleton Class A model with parameters given in (3.18) and (3.19). Environments with interference sources having similar rates, channel and emission statistics would also have a Middleton Class A statistics. To see this, assume that $\lambda_m = \lambda$, $\mu_m = \mu$, $\gamma(d_m) = \gamma$ and $\mathbb{E}\{\mathbf{g}_m^2\mathbf{B}_m^2\} = k \ \forall m \in 1, \cdots, M$. Substituting these values into (3.18) and (3.19) I get, for each interference source $m \in \{1, \cdots, M\}$,

$$A_m = \lambda\mu \qquad \Omega_m = \tfrac{\lambda\mu\times\gamma k}{2} \ . \tag{3.23}$$

Thus, the total interference $\mathbf{\Psi}$ is the sum of M independent Class A distributed random variables with parameters given by (3.23). Consequently, $\mathbf{\Psi}$ is also

Class A distributed with the following parameters [61]

$$A_\Psi = MA_m \qquad \Omega_\Psi = \Omega_m \ . \tag{3.24}$$

The variance of the noise is also multiplied by M. The assumption that $\gamma(d_m)$ is independent of d_m is especially valid in lower frequency PLC networks (order of 100kHz) since for wired sources

$$\gamma\left(d_m\right) = e^{-\alpha(f_0)d_m} = e^{-\left(a_0 + a_1 f_0^k\right)d_m} \approx 1$$

where the last equality follows by substituting some measured values of the given parameters: $a_0 = 0$, $a_1 = 7.8 \times 10^{-10}$, $f_0 = 100$kHz, $k = 1$ and d_m having typical PLC network dimensions ($20m \leq d_m \leq 500m$) [105]. In this range the transmission line effects are negligible and lumped discrete models can be used. These cases are summarized in Table 3.1.

3.1.7 Simulation and Experimental Results

I verify our derived models by using Monte-Carlo simulations of the system given in Figure 3.1. In particular, for each interferer source m I choose a *rate* λ_m and distance d_m such that $\lambda_m \sim \mathcal{U}\left(\lambda_{\min}, \lambda_{\max}\right)$ and $d_m \sim \mathcal{U}\left(d_{\min}, d_{\max}\right)$. I choose $\lambda_{\min} = 50$/sec and $\lambda_{\max} = 1000$/sec based on empirical measurements [50, 104]. Also, I choose to simulate a medium-sized PLC network with $d_{\min} = 50$m, $d_{\max} = 500$m, and $\alpha_0 = 10^{-4}$ with the number of interference sources $M = 5, 15$. The mean impulse duration $\mathbb{E}\left\{\mathbf{T}_m^E\right\}$ was chosen to be 150μsec $\forall m$ [104]. The accuracy of the statistical models is established by comparing the empirical tail probabilities based on the Monte-Carlo

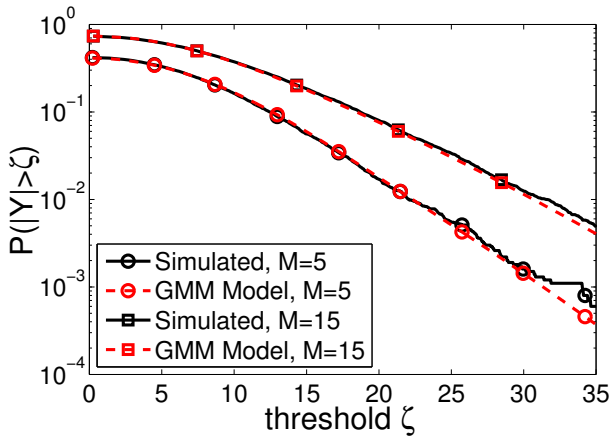


Figure 3.3: For M different interfering sources, empirical tail probabilities from Monte-Carlo simulations and the predicted tail probabilities from the Gaussian mixture model given in (3.22) are shown. In both cases, the curves match exactly for a wide range of ζ values.

simulated data and analytical tail probabilities predicted by our derived models. The tail probability characterizes the impulsiveness of a given distribution and is given by $\mathbb{P}(\|Y\| > y)$. The comparison between the two tail probabilities for the general cases where the number of interference sources is 5 and 15 is given in Figure 3.3. The empirical tail probability curve and the model predicted tail probability curve are exact matches with little deviation at the higher amplitudes due to the limited number of data points generated in that range. Moreover, the curves corresponding to the case with 15 interference sources is higher than that of 5 sources because the variance (power) is higher for the former. On the other hand, Figure 3.4 shows the tail probabilities for the homogeneous network described in Section 3.1.6. This network can be an appropriate approximation for low-frequency PLC networks and results in

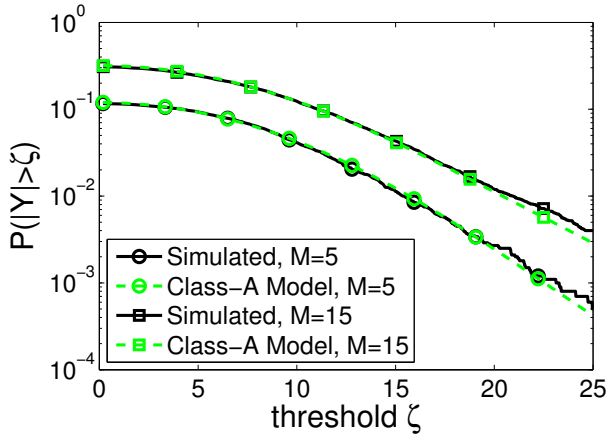


Figure 3.4: For M homogeneous interfering sources with similar statistics, empirical tail probabilities from Monte-Carlo simulations and the predicted tail probabilities from the Middleton Class A model given in (3.24) are shown. In both cases, the curves match exactly for a wide range of ζ values.

interference that is Middleton Class A distributed with parameters given in (3.24). Again, it can be seen that there is a good fit between the simulated data and the derived model.

In order to validate the above model, I captured real PLC network interference samples in an apartment building in Austin, TX. The noise was sampled in the $45-90$ kHz band at 1MSample/sec. I used the EM algorithm to fit the gathered data in chunks of 14 ms to the proposed models. The results, given by the tail probabilities, are shown in Figure 3.5. The Gaussian mixture model provides the best fit in accordance with our derived model. The Class A model does not fit well in this particular case indicating that the interference sources had different emission properties. As expected, the Gaussian model provided the worst fit because it does not take into consideration the heavy

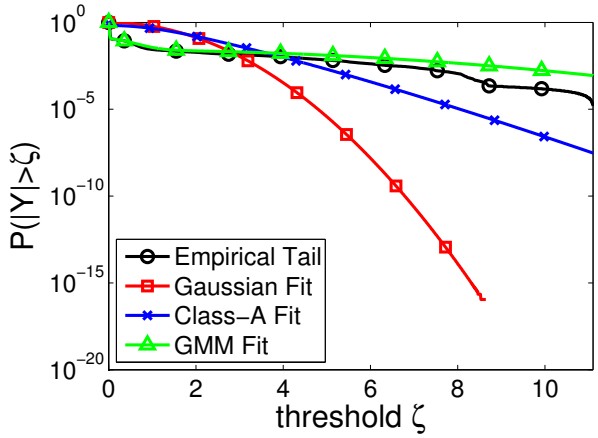


Figure 3.5: Comparison of tail probabilities obtained from measured data samples and by the Gaussian distribution, Middleton Class A distribution, and Gaussian mixture model. The Gaussian mixture model provides the best fit among the derived models.

tails of the interference distribution.

3.2 Cyclostationary Modeling of Narrowband PLC Interference

A Smart Grid intelligently monitors and controls energy flows in an electric grid. Having up-to-date distributed readings of grid conditions helps utilities efficiently scale generation up or down to meet demand. Narrowband powerline communication (PLC) systems can provide these up-to-date readings from subscribers to the local utility over existing power lines. While the interference in broadband PLC systems is dominated by asynchronous impulsive noise discussed in Section 3.1, a key challenge in narrowband PLC systems is overcoming additive non-Gaussian interference that exhibits strong

cyclic temporal dependence due to the non-linear coupling with the AC mains and various electrical devices employing switching power supplies. In this section, I propose to use a cyclostationary model for the dominant component of additive non-Gaussian noise. The key contributions are (1) fitting measured data from outdoor narrowband PLC system field trials to a cyclostationary model, and (2) developing a cyclostationary noise generation model that fits measured data.

3.2.1 Introduction

The increasing energy demands of the future necessitate non-traditional energy generation and management techniques. The concept of the *Smart Grid* addresses this issue by intelligently monitoring and controlling energy flows in the electric grid. A vital part of the Smart Grid revolves around providing reliable communication links between various agents in the network. A strong candidate for such a role is powerline communication (PLC) [33]. PLC technologies such as the TURTLE and TWACS have been in use by electric utilities for remote metering applications for two decades [33]. However, new Smart Grid applications demand much higher data rates than the one provided by those early PLC technologies. As a result, there has been a lot of interest in developing what is called high data rate *narrowband* ($3 - 500$ kHz) PLC systems for remote metering and load control. Examples of such systems are the ongoing standards such as ITU-T G.hnem and IEEE 1901.2 and the proprietary PRIME and G3. These systems employ OFDM modulation to provide

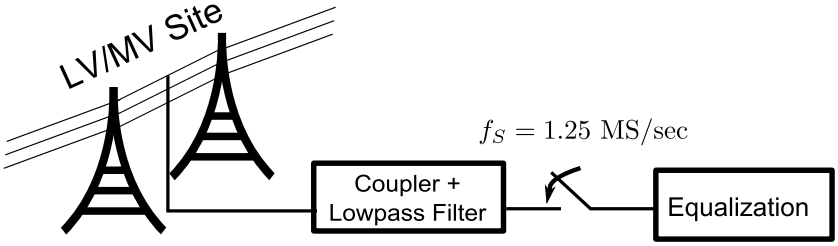


Figure 3.6: The measurement setup listens to the powerline communication band on low-voltage (LV) and medium-voltage (MV) lines, and samples the noise traces at 1.25 MS/sec.

data rates up to hundreds of kilobits per second.

The attractive aspect of PLC is the possible deployment over the existing power grid, thereby saving the cost of a new infrastructure. The downside is that this infrastructure, originally designed for one-way power transfer, is a hostile environment for communication systems. Time-varying non-Gaussian noise and time-varying frequency selective channels are the two primary impairments affecting reliable PLC [29, 33]. This section focuses on noise modeling for narrowband PLC systems. I refer the reader to [29] for PLC channel models.

There has been significant interest in characterizing PLC noise due to its impact on communication performance. Various noise models have been proposed to capture the noise characteristics in PLC environments in frequency ranges up to 20 MHz. Generally, PLC noise can be viewed as an aggregation of various types of noise [29, 72, 104]. Many properties of PLC noise have been studied empirically in [104]. However, these studies focus on the noise in the $0.2 - 20$ MHz range and thus are more applicable for broadband PLC sys-

tems. Less work has been done on characterizing narrowband PLC noise. An exception is the periodic noise model proposed in [81] for the very low frequency PLC and the *cyclostationary Gaussian* proposed in [52] that captures the temporal cyclic behavior that dominates narrowband PLC noise. However, this model ignores the time-varying spectral behavior of the noise which limits its applicability to narrow single carrier systems, making it inappropriate for OFDM systems. This spectral variation results from the noise being the superposition of various noise processes with different generation mechanisms (such as homes, heavy industry). Furthermore, the measurements used in [52] were taken in indoor environments and don't generalize readily to outdoor environments such as the ones employed by utilities.

3.2.2 Contribution

In this section, I present measurements results from a low voltage site. Then, I propose a passband cyclostationary noise model for narrowband PLC that accounts for both the time and frequency properties of the measured noise. The proposed model is computationally tractable and can be exploited by the PLC modem for link adaptation. This work has been done in the context of the IEEE P1901.2 standardization effort [23].

3.2.3 Measurement Setup

The measurement setup is shown in Figure 3.6. The analog to digital converter (ADC) connects to a low voltage or medium voltage power line

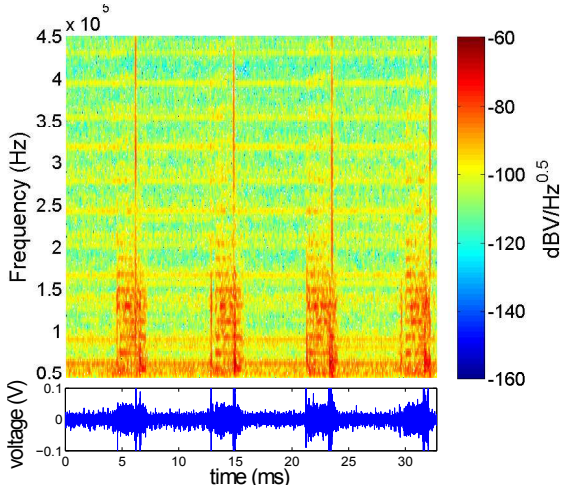


Figure 3.7: Spectrogram of a noise trace at a low voltage site [23]. The noise displays the cyclostationary features both in time and frequency.

through a coupler and listens to the PLC environment under signal silence. Since I are interested in narrowband PLC noise, a low pass filter with a cut-off frequency of around 500 kHz is utilized. The output of this filter is sampled at a sampling rate $f_S = 1.25$ MS/sec. Before analyzing the data, I remove the effect of the spectral shape of the acquisition equipment through equalization.

3.2.4 Data Analysis

Communication systems models need to capture both the temporal and spectral properties of the noise. A commonly used technique for non-stationary signal analysis is the *Short-Time Fourier Transform* (STFT) [76]. The resulting spectrogram (magnitude of the STFT) of a noise trace collected at a low voltage site is given in Figure 3.7. This noise exhibits strong cyclostationary features in time and frequency domain with period $T = T_{AC}/2 \approx 8.3$ ms. In

addition, there is a higher concentration of noise power in the lower frequency band with broadband impulses occurring every T and some weaker narrowband interference. A complete analysis of 22 low voltage and medium voltage sites are given in [23].

3.2.5 Cyclostationary Gaussian Model

The *Cyclostationary Gaussian Model* (CGM) is a cyclostationary model proposed in [52] to model the dominant noise in narrowband PLC systems. According to this model, the passband noise samples are modeled as zero-mean Gaussian random variables with a periodic time-varying variance $\sigma^2[k]$ of period N; i.e.

$$s[k] \sim \mathcal{N}\left(0, \sigma^2[k]\right), \quad \sigma^2[k] = \sigma^2[k + lN] \tag{3.25}$$

where k is the time index and $l \in \mathbb{Z}$. The period $N = Tf_S$ where f_S is the sampling frequency[1]. The variance $\sigma^2[k]$ is modeled as a sum of L sinusoids with $3L$ parameters. The resulting noise process $s[k]$ is cyclostationary with autocorrelation given by

$$r_s\left[k, \tau\right] = \mathbb{E}\left\{n[k]n[k+\tau]\right\} = \sigma^2[k]\delta[\tau]. \tag{3.26}$$

As expected, $r_s[k, \tau] = r_s[k + N, \tau]$. Due to $\delta[\tau]$, the spectrum of this process is white in frequency with time-varying power. As a result, the CGM shapes the resulting $s[k]$ with an LTI filter $h[k]$ to produce a decaying spectral profile

[1] f_S is assumed to be aligned with T to result in $N \in \mathbb{N}$.

independent of time. The LTI filter is chosen to fit the spectral shape of the background noise typically assumed to be exponentially decaying [29]. The autocorrelation of the resulting process $n[k]$ is given by

$$r_n[k,\tau] = \sum_m h[m]\sigma^2[k-m]h[\tau+m]. \tag{3.27}$$

While still periodic, the resulting correlation is coupled with $\sigma^2[k-m]$ and the resulting spectrum no longer corresponds to the shaping filter $h[k]$. Furthermore, there is no physical basis for choosing the sinusoid as the parametric form for $\sigma^2[k]$. This leads to a huge expansion in the parameter space, particularly if the noise envelope has sharp transitions as seen in Figure 3.7, requiring large amount of data and complexity for parameter estimation ($50-100$ AC cycles [52]).

3.2.6 Proposed Cyclostationary Model

The CGM models the noise process as an excitation of an LTI system $h[k]$ by a cyclostationary input $n[k]$ given in (3.25). While accurate for background noise, a single LTI system $h[k]$ doesn't capture the time variation of the spectral content shown in Figure 3.7 which represent the aggregation of various physical phenomena. This mismatch in the spectral domain makes this noise model inappropriate for modern PLC standards that employ OFDM [29]. Given the limited applicability of CGM to OFDM systems, I propose a noise model for narrowband PLC that takes into account both the spectral and temporal properties of the noise.

3.2.6.1 Spectral Modeling

Figure 3.7 shows that the noise spectral content has three distinct regions in each period T where the spectrum has similar shape corresponding to a specific generating physical phenomena: a low power background noise region ($0-5$ ms in Figure 3.7), a high power interference region ($5-7$ ms in Figure 3.7), and a broadband impulse of duration ≈ 0.3 ms. In general, a given period of duration T can be divided into M intervals $\mathcal{R}_1, \cdots, \mathcal{R}_M$ where the noise spectral shape remains unchanged (M is between 2 and 4 [23]). If I assume that the noise is stationary in each interval $\mathcal{R}_i$, then I can model the noise in that interval as a response of an LTI filter $h_i[k]$ to a stationary input $s[k]$. Accordingly, the noise can be modeled as the response of a *linear periodically time-varying* (LPTV) system $h[k,\tau]$ to a stationary input $s[k]$ where

$$h[k,\tau] = \sum_{i=1}^{M} h_i[\tau]\mathbf{1}_{k\in\mathcal{R}_i}, \quad 0 \le k \le N-1 \tag{3.28}$$

and $h[k+lN,\tau] = h[k,\tau]$ where N is the discrete period corresponding to half the AC-cycle T, $l \in \mathbb{Z}$, and $\mathbf{1}_{\mathcal{A}}$ is the indicator function ($\mathbf{1}_{\mathcal{A}} = 1$ if $\mathcal{A}$, 0 otherwise). As a result, the noise $n[k]$ is given by

$$n[k] = \sum_{\tau} h[k,\tau]s[\tau] = \sum_{i=1}^{M} \mathbf{1}_{(k \bmod N)\in\mathcal{R}_i} \sum_{\tau} h_i[\tau]s[\tau]. \tag{3.29}$$

This can be interpreted as sequential filtering of the stationary input $s[k]$ by a sequence of LTI filters $h_i[k]$ (See Figure3.8). The LPTV system approach is further motivated by [21] where the indoor PLC channel response was shown

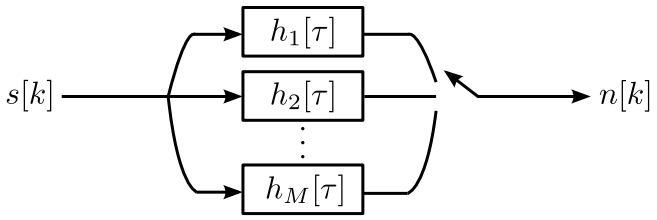


Figure 3.8: Noise generation model: $n[k]$ is the result of sequential filtering of stationary input $s[k]$ by a sequence of LTI filters $h_i[\tau]$.

to be well approximated by a LPTV filter consisting of a sequence of time invariant filters.

3.2.6.2 First-Order Statistics of the Noise Samples

The LPTV filtering operation models the second order statistics of the cyclostationary noise. In this section, I examine the first-order statistics of the cyclostationary noise $n[k]$ to determine the appropriate excitation stationary process $s[k]$ ($n[k]$ is a weighted sum of $s[k]$ samples). For a cyclostationary process,

$$p_k(z) = p_{k+lN}(z), \quad l \in \mathbb{Z} \tag{3.30}$$

where $p_k(z)$ is the pdf of the noise sample $n[k]$. As a result, the pdf $p_k(z)$ can be estimated from the pdf of the subsampled process $n_k[l] = n[k + lN]$. Figure 3.9 indicates that the normal distribution can be a good fit for the subsampled sequences $n_k[l]$. The Lilliefors test for normality over a noise trace of 12 periods shows that 95% of the submsampled sequences $n_k[l]$ fit the normal distribution at a significance level $\alpha = 0.01$. Since filtering a Gaussian process by a linear system produces another Gaussian process, $s[k]$ can be

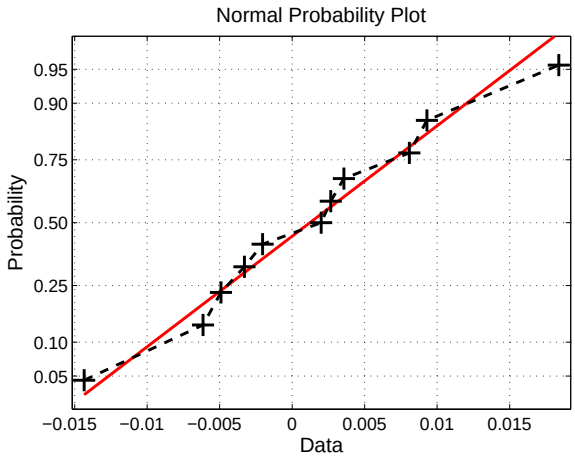


Figure 3.9: The normal plot for 12 samples from the subsampled $n_{k_0}[l]$ for some k_0. The closeness of the data points to the red line indicates that the samples follow a normal distribution.

modeled as a Guassian process. To simplify the estimation of the shaping filters $h_i[k]$, I make $s[k]$ a unit power Gaussian white noise.

3.2.6.3 Parameter Estimation

The proposed model is parametrized by the number of stationary regions M, the region intervals $\{\mathcal{R}_i : 1 \leq i \leq M\}$, and the LTI filters $\{h_i[k]\}_{1 \leq i \leq M}$. The number of stationary regions M and the region boundaries can be inferred by visually inspecting the spectrogram such as the one in Figure 3.7. Furthermore, the stationary assumption during each interval $\mathcal{R}_i$ allows for an efficient automated region detection in the time domain that can be implemented on an PLC receiver. In particular, under the assumption that each LTI filter $h_i[k]$ has a different power $\|\mathbf{h}_i\|^2$ (as is typically the case [23]), each noise sample

$n[k]$ will have a power given by

$$\mathbb{E}\left\{n^2[k]\right\} = \|\mathbf{h}_i\|^2, \qquad k \in \mathcal{R}_i \tag{3.31}$$

due to the stationarity assumption. This means that noise samples within each region have equal powers. As a result, a simple thresholding scheme might be adopted to differentiate regions in the time-domain. Furthermore, a PLC modem can set the thresholds γ_i to correspond to its adaptive coding and modulation thresholds; thus estimating only the noise parameters that are relevant to the communication performance.

The LTI filters $\{h_i[k] : 1 \leq i \leq M\}$ are spectrum shaping filters. Designing these filters requires a spectrum estimate for each region $\mathcal{R}_i$. Parametric and non-parametric techniques for spectral estimation are discussed in [48]. The trade-off between using either method is estimation accuracy vs. generalization error. Parametric models produce more accurate estimates under the correct model assumptions but suffer under model mismatch. On the other hand, non-parametric models generalize well but suffer from an increased noise floor. In narrowband PLC, the spectral shapes vary significantly between sites and the time of the day and may include narrowband interferers [23]. As a result, non-parametric models are more appropriate for designing robust PLC systems for field deployment. Given an estimate of the spectrum $\hat{S}_i(\omega)$ during $\mathcal{R}_i$, an estimate of the autocorrelation sequence $\hat{r}_i[\tau]$ during that same interval can be obtained by taking its IDFT. This sequence can be then used to design the appropriate spectrum shaping filter $h_i[k]$ [48]. In addition, frequency

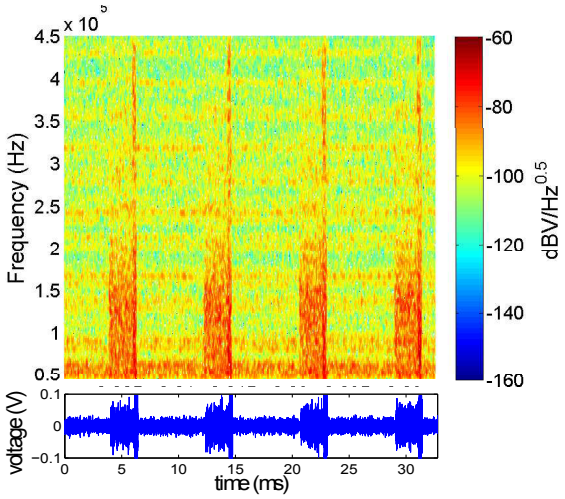


Figure 3.10: The spectrogram of the fitted model: a close match to the spectrogram of the PLC noise given in Figure 3.7.

domain filtering using FFT can be applied using the spectral estimate $\hat{S}_i(\omega)$ followed by an IDFT operation.

3.2.7 Model Fitting

The application of the proposed model to narrowband PLC, in particular OFDM, depends on its accuracy in modeling the spectral properties of the PLC noise. I apply the proposed modeling procedure to the data displayed in Figure 3.7. By visual inspection, I determine $M = 3$ and the intervals $\mathcal{R}_1, \mathcal{R}_2$, and $\mathcal{R}_3$ corresponding to the regions described in Section 3.2.4. The corresponding spectral estimates $\hat{S}_1(\omega), \hat{S}_2(\omega)$, and $\hat{S}_3(\omega)$ are estimated using the Welch's method [48]. Applying frequency domain filtering to a unit power AWGN noise, the spectrogram for the generated noise is given in Figure 3.10. As shown in Figure 3.10, the fitted model generates noise samples

whose spectral and temporal traces resembles closely that of the original data.

3.3 Empirical Modeling of ISM-Band Wireless Interference

In this section, I propose two empirical models for characterizing interference in the ISM band: the *Gaussian mixture* (GM) model, discussed previously, and the *Gaussian hidden Markov* (GHMM) model. Each of these models captures a specific feature of the impulsive noise observed in practice: the GM model captures the marginal pdfs of the impulsive noise while the GHMM extends it to capture local dependencies in bursty impulsive noise samples. I briefly review the GM model and its interpretation under a latent-variable model as discussed in Section 2.5.1. This allows us to relate it to the GHMM in Section 3.3.2.

3.3.1 The Gaussian Mixture Models - Impulsive Noise

The GM distribution is a generalization of the Gaussian distribution that allows for multiple modes and heavier tails. The latter is what makes the GM distribution suitable for impulsive noise modeling since large amplitudes (the pdf tails) are more likely in impulsive environments than in AWGN. A GM distribution of a random variable n with zero mean components is given by

$$p\left(n\right)=\sum_{k=0}^{K-1}\pi^{(k)}\cdot\mathcal{N}\left(n;0,\gamma_{k}\right) \tag{3.32}$$

where $\mathcal{N}(n; 0, \gamma_k)$ denotes a complex Gaussian pdf with zero mean and variance γ_k; and $\pi^{(k)}$ is the probability of the k-th Gaussian component. Typically, the component with the smallest variance represents the Gaussian background noise g and the total noise can be decomposed as the sum of background and impulsive noise given as

$$n = g + i \tag{3.33}$$

where i is the impulsive noise component. If I let $g \sim \mathcal{N}(0, \gamma_0)$, then the pdf of the impulsive noise is given by

$$p(i) = \pi^{(0)}\delta(i) + \sum_{k=1}^{K-1} \pi^{(k)} \cdot \mathcal{N}(i; 0, \gamma_k), \tag{3.34}$$

i.e. with probability $\pi^{(0)}$ only background noise is present and with probability $1-\pi^{(0)}$ I have a non-zero impulsive noise on top of the background noise. Here, the variances γ_k indicate the power of the impulse component rather than the total variance of the noise n as in (3.32) (γ_k in (3.32) equals $\gamma_0 + \gamma_k$ in (3.34)). The mixing vector $\boldsymbol{\pi} = [\pi^{(0)} \cdots \pi^{(K-1)}]$ can be interpreted as a probability mass function of a latent random variable z that selects the mixture component used to generate a sample of n; *i.e.* $P(z = k) = \pi^{(j)}, \forall k \in \{0, \cdots, K-1\}$. Given z,

$$i|\{z = k\} \quad \sim \mathcal{N}(0, \gamma_k) \tag{3.35}$$

where $\gamma_0 = 0$. Middleton class-A distribution is a special case of the GM distribution with infinite number of Gaussian components [69]. In practice, these are truncated to two or three components and the finite GM pdf given in (3.32) provides an accurate approximation [97]. By choosing an appropriate

model order, GM models provide flexibility in modeling general distributions [64].

3.3.2 The Gaussian Hidden Markov Models - Bursty Noise

In many scenarios, such as in PLC systems, the noise is not only impulsive but also bursty; *i.e.*, impulses arrive in bursts [104]. As a result, the noise samples are no longer independent from each other as in the GM case. This dependence can be captured by hidden Markov models (HMM) with Gaussian emission probabilities [32]. In [104], the authors show that a Markov chain with 7 states is able to accurately reproduce the burst durations observed in broadband PLC.

The Gaussian HMM (GHMM) can be viewed as a generalization of the GM model to the case where the latent variable at time t, z_t, depends only the previous latent variable z_{t-1}. The temporal dynamics of this Markov chain are determined by the state transition matrix $\mathbf{T}$ defined as

$$[\mathbf{T}]_{i,j} = P(z_t = j | z_{t-1} = i) \quad \forall i, j \in \{0, \cdots, K-1\} \tag{3.36}$$

where I assumed that the Markov chain is homogeneous (stationary). Note that the marginal distribution of the noise samples under this model is GM with pdf given by (3.34) with the mixing vector $\boldsymbol{\pi}$ given by the solution of $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{T}$. Under this model, the mean duration of a bursty impulse noise from state k is given by

$$\mathbb{E}\{\tau_k\} = \frac{1}{1 - [\mathbf{T}]_{k,k}} \tag{3.37}$$

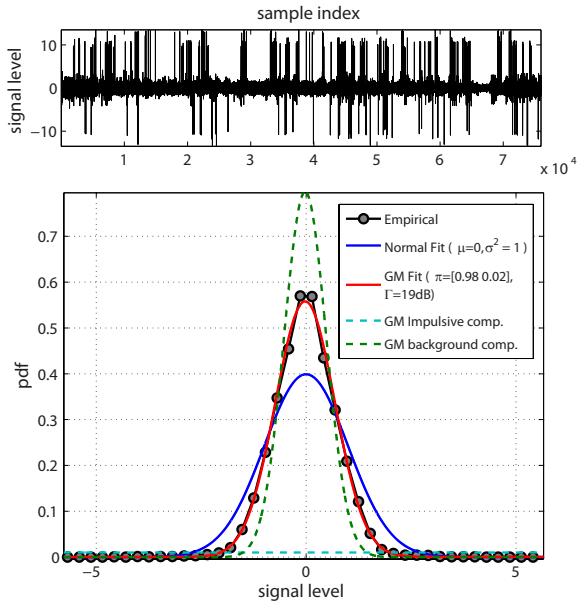


Figure 3.11: A pdf fit of a noise trace collected from a receiver embedded in a laptop: the GM model with 2 components provides a significantly better fit than the normal fit.

where τ_k is the duration that the chain will persist in state k [32]. Given the value of the latent variable z_t the noise value at time t is selected according to (3.35).

3.3.3 Fitting Empirical Data

Assuming the noise statistics are slowly varying, the parameters $\boldsymbol{\pi}$, $\{\gamma_k\}_{k=0}^{K-1}$, and $\mathbf{T}$ can be estimated using the expectation-maximization (EM) algorithm [12] during the quiet time when there is no signal transmission. As an illustrative example, I fit a noise trace collected from a laptop and shown in Figure 3.11 into a 2-state GHMM. Using the EM algorithm, I find that

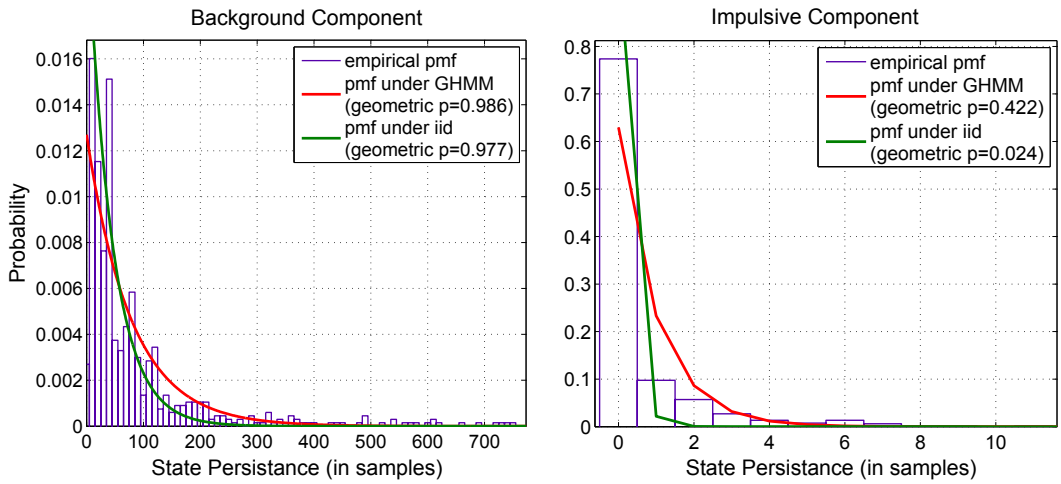


Figure 3.12: The distribution of the persistence time (time to remain in a state once its entered) of each state and its fit under the iid GM model and the GHMM: the GHMM provides a better fit especially for the impulsive component duration.

$\mathbf{T} = \begin{bmatrix} 0.422 & 0.578 \\ 0.013 & 0.986 \end{bmatrix}$ and $\Gamma = \gamma_1/\gamma_0 = 17$dB with a marginal given by a GM model with $\boldsymbol{\pi} = \begin{bmatrix} 0.98 & 0.02 \end{bmatrix}$. Figure 3.11 shows that GM model, unlike the normal model, provides a close fit to the empirical pdf of the collected data (fitted using a kernel density estimator using a Gaussian kernel). Furthermore, Figure 3.12 shows that there is temporal dependence between the noise samples (especially in the impulsive component) that is captured by the GHMM model.

3.4 Conclusion

In this chapter, I derive statistical-physical models for uncoordinated interference in PLC networks. In particular, I show and verify by simulation that the interference in homogeneous PLC networks follows the Middleton

class-A model while the interference in dense PLC networks follows the more general Gaussian mixture model. This is inline with experimental studies reported in the literature [26, 104] and similar to models derived for wireless networks [41, 42]. In addition, I show that a cyclostationary model is appropriate for modeling the periodicity exhibited by interference in narrowband PLC and that a Gaussian HMM model captures the statistics of bursty interference present in some wireless platforms operating in the 2.4GHz ISM band.

In the following chapters, I use these models in the design of OFDM receivers corrupted by interference. In particular, I focus on the Gaussian mixture and Gaussian hidden Markov models. These models provide natural priors on the interference. By exploiting this prior knowledge, the proposed receiver are able to improve the communication performance by tens of dBs.

Chapter 4

EM-Based OFDM Receiver in Gaussian Mixture Interference

Thus far, I have argued that interference cannot be completely eliminated in modern OFDM systems such as cellular LTE and PLC networks. Furthermore, Section 2.3 and Chapter 3 show that uncoordinated interference in PLC and wireless networks can be modeled using statistical-physical and empirical models. However, as discussed in Section 2.1, even though OFDM modulation provides resilience to impulsive noise due to its code diversity, OFDM receivers are commonly designed assuming that the noise and interference is AWGN, mainly due to the computational tractability provided by independent decoding across subcarriers. This leads to suboptimal performance due to the dependence in noise statistics across subcarriers resulting from the DFT operation applied during receiver processing. Due to this dependence, optimal detection of OFDM symbols becomes prohibitive due to its exponential complexity. In this chapter, I consider the design of a practical class of OFDM receivers that are constrained to perform independent detection on each subcarrier. I propose an EM based low-complexity iterative decoding algorithm for OFDM systems in impulsive noise environments that preserves the independent decoding across subcarriers. I validate its performance under

typical impulsive noise conditions in wireless and powerline platforms. The proposed method achieves a gain between 2-7dB over the conventional OFDM receiver depending on the SNR range.

4.1 Introduction

Communication transceivers in powerline communication (PLC) and wireless networks suffer from uncoordinated interference from other users and non-communication sources such as microwave ovens and switching power supplies. The typical additive white Gaussian noise (AWGN) assumption is inadequate for capturing the statistical properties of such interference and leads to suboptimal receivers. Different statistical models for impulsive noise have been discussed in Chapter 3 in order to help in designing and analyzing the performance of receivers in the impulsive noise channels.

The advantages of OFDM have been discussed in Section 1.2. Due to these advantages, OFDM modulation has been adopted in many modern wireless communication standards, such as IEEE802.11n and LTE, and recent PLC standards, such as PRIME and G3. As discussed in Section 2.4, the design of OFDM receivers in impulsive noise can be classified into two subcategories according to the assumed impulsive noise model. On one hand, non-parametric techniques do not assume any particular model for impulsive noise and treat it as a sparse vector. Example of such a technique is a compressed-sensing approach has been proposed in [17], while a more general Sparse Bayesian Learning approach was given in [57]. On the other hand, parametric models assume

a given impulsive noise model and design the receiver based on its statistical properties. In general, parametric approaches outperform non-parametric models if the assumed model provides a good match for the underlying impulsive noise ([57], Chapter 5). This is typically the case in slowly varying environments with quiet period that can be used for parameter training.

4.2 Contribution

In this chapter, using the GM model for uncoordinated interference, I propose a parametric EM-based low-complexity iterative decoding algorithm for OFDM systems in impulsive noise environments that preserves independent decoding across subcarriers.

4.3 System Model

I consider the simplified OFDM system, described by the discrete-time baseband model

$$\mathbf{y} = \underbrace{\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}}_{\mathbf{u}} + \mathbf{w} \tag{4.1}$$

where $\mathbf{y} = [y_1 \cdots y_N]^T$ is the received signal with N being the FFT block length (number of subcarriers), ρ is signal power, $\mathbf{x}$ is the $N \times 1$ frequency-domain transmitted symbols, and $\mathbf{w} = [w_1 \cdots w_N]^T$ is the $N \times 1$ additive noise vector. The channel matrix $\mathbf{H}$ is a circulant matrix whose elements are assumed to be known at the receiver. The FFT operation is represented by the $N \times N$ matrix $\mathbf{F}$ where $(\cdot)^*$ represents the Hermitian operator. The $N \times 1$

vector $\mathbf{u}$ is the time-domain transmitted OFDM signal. The noise is assumed to be temporally independent and identically distributed (*i.i.d.*) Gaussian mixture random vector. Thus, the probability density function (*pdf*) of $\mathbf{w}$ is the product of single variable GM model pdf given by

$$p(\mathbf{w}) = \prod_{i=1}^{N} \sum_{j=1}^{K} \pi_j \mathcal{N}\left(w_i; 0, \sigma_j^2\right). \tag{4.2}$$

4.4 Optimal OFDM Detection in Impulsive Noise

The problem of detecting an OFDM symbol for the system model given in (4.1) can be formulated as

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}) = \arg\max_{\mathbf{x}} p_{\mathbf{w}}(\mathbf{y} - \sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}). \tag{4.3}$$

In (4.3), each of the components of $\mathbf{w}$ depends on $y_j - \sqrt{\rho}\left[\mathbf{H}\mathbf{F}^*\mathbf{x}\right]_j$ which is a function of all components of $\mathbf{x}$. In addition, there is no efficient code representation for $\mathbf{F}$ which would reduce the decoding complexity. As a result, an exhaustive search would be required to solve this problem. Conventional OFDM receivers, designed under the Gaussian noise assumption, circumvent this problem by computing the following statistic

$$\Psi = \mathbf{F}\mathbf{y} = \sqrt{\rho}\mathsf{H} \circ \mathbf{x} + \underbrace{\mathbf{F}\mathbf{w}}_{\mathbf{z}}. \tag{4.4}$$

where H is the frequency domain channel. When $\mathbf{w}$ is Gaussian, the transformed noise $\mathbf{z}$ has a product form *pdf* across subcarriers because $\mathbf{F}$ is unitary and preserves the Gaussian statistics of $\mathbf{w}$ and the independence between the

noise vector samples in the Fourier domain. As a result, Ψ is a sufficient statistic and decoding can be performed independently across subcarriers. However, for the noise model in (4.2), the transformed noise $\mathbf{z}$ has dependent components which means that detection across subcarriers can not be decoupled as in the Gaussian case. This leads to the same exhaustive search as in (4.3).

4.5 Low Complexity Suboptimal Decoders

Due to the dependance of noise samples in the Fourier domain, optimal detection can not be performed independently across subcarriers and has an exponential complexity in the number of subcarriers N (ranging from 64 to 1024 for modern communication systems). This makes optimal detection based on (4.3) impractical on current computational platforms. In addition, many communication system assume independent decoding across subcarriers. As a result, it desirable to design algorithms that will improve performance under such a constraint. Two observations, employed by [45] to simplify the problem, are: 1) the noise $\mathbf{w}$ is *i.i.d.* in time, and 2) the time-domain signal $\mathbf{u} = [u_1 \cdots u_N]^T$ in (4.1) can be approximated, due to the Central Limit Theorem, as being *i.i.d.* in time and $u_j \sim \mathcal{N}_c\left(0, \|\mathbf{h}\|^2\rho\right), \forall j$ where $\mathbf{h}$ is the known channel vector. [45] then proceeds to find the MMSE estimate, $\hat{\mathbf{u}}$, of $\mathbf{u}$ with NSI and without NSI, followed by hard detection on $\mathbf{F}\hat{\mathbf{u}}$. Since the proof is not explicitly given in [45], I provide it here for completeness.

4.5.1 MMSE Estimation with NSI

When NSI is available, the noise at time j is Gaussian with variance $\sigma^2_{s_j}$. The NSI is given by vector $\mathbf{s} = [s_1 \cdots s_N]^T$ where s_j represents the state of the noise at the time instance j (see Section 2.5.1). Let Λ be a matrix function of $\mathbf{s}$ given by

$$\Lambda(\mathbf{s}) = \text{diag}\left\{1/\sigma_{s_1}, \cdots, 1/\sigma_{s_N}\right\}. \tag{4.5}$$

Multiplying (4.1) by $\Lambda(\mathbf{s})$, I obtain

$$\Lambda(\mathbf{s})\,\mathbf{y} = \Lambda(\mathbf{s}) \underbrace{\mathbf{u}}_{\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}} + \underbrace{\Lambda(\mathbf{s})\,\mathbf{w}}_{\mathbf{n}} |\mathbf{s} \tag{4.6}$$

where $\mathbf{n}$ is now a Gaussian vector with identity covariance matrix. However, independent detection across subcarriers would introduce intersymbol interference (ISI) in the frequency domain since $\mathbf{F}\Lambda(\mathbf{s})\,\mathbf{F}^* \neq \mathbf{I}_N$. Since $\mathbf{u}$ and $\mathbf{n}$ are Gaussian, the MMSE estimate of $\mathbf{u}$ is also the Linear MMSE estimate given by

$$\hat{\mathbf{u}}(\mathbf{y}, \mathbf{s}) = \text{diag}\left\{\frac{\|\mathbf{h}\|^2\rho}{\|\mathbf{h}\|^2\rho + \sigma^2_{s_1}}, \cdots, \frac{\|\mathbf{h}\|^2\rho}{\|\mathbf{h}\|^2\rho + \sigma^2_{s_N}}\right\}\mathbf{y}. \tag{4.7}$$

At any time instant j, (4.7) multiplies the observation by $\frac{\|\mathbf{h}\|^2\rho}{\|\mathbf{h}\|^2\rho+\sigma^2_{s_j}}$. This scaling reflects the reliability of the received sample based on the noise state it was received under. The implementation complexity of this estimator is low, however the assumption of having NSI at the receiver does not hold in most practical scenarios.

4.5.2 MMSE Estimation without NSI

When NSI is not present at the receiver, (4.1) can not be normalized as in (4.6) and the resulting MMSE estimator $\hat{\mathbf{u}} = [\hat{u}_1 \cdots \hat{u}_N]^T$ of $\mathbf{u}$ is a nonlinear function of $\mathbf{y}$. It can be shown that the MMSE estimate is given by

$$\hat{u_j} = \frac{\mathrm{E}_s\left[\frac{\|\mathbf{h}\|^2\rho}{(\|\mathbf{h}\|^2\rho+\sigma_s^2)^2}\exp\left(-\frac{\|y_j\|^2}{\|\mathbf{h}\|^2\rho+\sigma_s^2}\right)\right]}{\mathrm{E}_s\left[\frac{1}{\|\mathbf{h}\|^2\rho+\sigma_s^2}\exp\left(-\frac{\|y_j\|^2}{\|\mathbf{h}\|^2\rho+\sigma_s^2}\right)\right]} \cdot y_j \tag{4.8}$$

where the index j is dropped from the expectation. The proof is given in Appendix 4.11.

4.6 The EM Algorithm

The EM algorithm is an iterative algorithm used to compute the ML estimate of a desired parameter $b \in \mathcal{B}$ given some observed data $y \in \mathcal{Y}$. In particular, it solves the following optimization

$$\hat{b} = \arg\max_{b\in\mathcal{B}} p\left(y|b\right) \tag{4.9}$$

where $p\left(y|b\right)$ is the conditional density of y given b. In order to achieve this, it treats this problem as incomplete data estimation problem where the missing data α simplifies the evaluation of $p\left(y,\alpha|b\right)$). The EM algorithm uses the likelihood function of the complete data in a two-step procedure as follows:

1. E-step: Compute $Q\left(b|\hat{b}^i\right) = \mathrm{E}_\alpha\left[\log f\left(y,\alpha|b\right)|y,\hat{b}^i\right]$
2. M-step: Solve $\hat{b}^{i+1} = \arg\max_{b\in\mathcal{B}} Q\left(b|\hat{b}^i\right)$

Given the right initial conditions, the estimate $\hat{b}^i$ will converge to a stationary point. In general, the solution of (4.9) can be obtained by an appropriate choice of the initial value. In communication systems, the EM algorithm has been widely applied to sequence and channel estimation problems. In [34], the authors give a detection-specific framework for applying EM to sequence estimation problems in communication systems.

4.7 Proposed EM-based Detection Algorithm

The ML detection rule of the transmitted vector $\mathbf{x}$ is given by

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p\left(\mathbf{y}|\mathbf{x}\right). \tag{4.10}$$

In Section 4.5.1, NSI reduced the complexity of the MMSE estimation from a non-linear function to a linear function of $\mathbf{y}$. This suggests that the latent vector of noise states $\mathbf{s}$ could be an appropriate choice for unobserved data in an EM-implementation. Thus, I choose $(\mathbf{y}, \mathbf{s})$ as our complete data and formulate the E-step accordingly. The likelihood of the complete data can be written as

$$p\left(\mathbf{y}, \mathbf{s}|\mathbf{x}\right) = p\left(\mathbf{y}|\mathbf{s}, \mathbf{x}\right) p\left(\mathbf{s}|\mathbf{x}\right) = p\left(\mathbf{y}|\mathbf{s}, \mathbf{x}\right) p\left(\mathbf{s}\right) \tag{4.11}$$

where the second equality follows from the fact that $\mathbf{x}$ and $\mathbf{s}$ are independent (transmission is not adapted to noise state). Since $p\left(\mathbf{s}\right)$ is not a function of $\mathbf{x}$, it will not have an effect on the M-step and can be ignored. Given that $\mathbf{y}$ is

Gaussian given $\mathbf{s}$ and $\mathbf{x}$, the E-step can be expressed as

$$\begin{aligned} Q\left(\mathbf{x}|\hat{\mathbf{x}}^i\right) &= \mathrm{E}_{\mathbf{s}}\left\{\log p\left(\mathbf{y}|\mathbf{s},\mathbf{x}\right)|\mathbf{y},\hat{\mathbf{x}}^i\right\} \\ &\overset{(1)}{=} \mathrm{E}_{\mathbf{s}}\left\{-\left(\mathbf{y}-\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}\right)^*\Lambda_{\mathbf{s}}^{-1}\left(\mathbf{y}-\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}\right)|\mathbf{y},\hat{\mathbf{x}}^i\right\} \\ &\overset{(2)}{=} -\left(\mathbf{y}-\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}\right)^*\mathrm{E}_{\mathbf{s}}\left\{\Lambda_{\mathbf{s}}^{-1}|\mathbf{y},\hat{\mathbf{x}}^i\right\}\left(\mathbf{y}-\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}\right) \end{aligned}$$

where $\Lambda_{\mathbf{s}} = \mathrm{diag}\left\{\sigma_{s_1}^2,\cdots,\sigma_{s_N}^2\right\}$ is the covariance matrix of $\mathbf{y}$ given $\mathbf{s}$ and $\mathbf{x}$. The term $\mathrm{E}_{\mathbf{s}}\left\{\Lambda_{\mathbf{s}}^{-1}|\mathbf{y},\hat{\mathbf{x}}^i\right\}$ is a diagonal matrix as well with diagonal entries $\frac{1}{\gamma_j^i}, \forall j \in \{1,\cdots,N\}$ given by

$$\frac{1}{\gamma_j^i} = \sum_{s_j=1}^{K}\frac{1}{\sigma_{s_j}^2}p\left(s_j|\mathbf{y},\hat{\mathbf{x}}^i\right) = \sum_{s_j=1}^{K}\frac{\pi_{s_j}}{\sigma_{s_j}^2}\frac{p\left(y_j|s_j,\hat{\mathbf{x}}^i\right)}{p\left(y_j|\hat{\mathbf{x}}^i\right)} \tag{4.12}$$

where the second equality follows from the application of Bayes rule and substituting for the corresponding probabilities. The term $p\left(y_j|\hat{\mathbf{x}}^i\right)$ is a constant with respect to s_j and can be computed as the normalization constant for the distribution $p\left(s_j|\mathbf{y},\hat{\mathbf{x}}^i\right)$ as follows

$$p\left(y_j|\hat{\mathbf{x}}^i\right) = \sum_{s_j=1}^{K}\pi_{s_j}p\left(y_j|s_j,\hat{\mathbf{x}}^i\right).$$

As a result, the only term that requires non-linear computation is $p\left(y_j|s_j,\hat{\mathbf{x}}^i\right) = \frac{1}{\pi\sigma_{s_j}^2}e^{-|y_j-\sqrt{\rho}[\mathbf{F}^*\hat{\mathbf{x}}^i]_j|^2/\sigma_{s_j}^2}$ which can be implemented using a look-up table. Let $\Gamma_{\mathbf{y},\hat{\mathbf{x}}^i} = \mathrm{diag}\{\gamma_1^i,\cdots,\gamma_N^i\}$, then the M-step can be written as

$$\hat{\mathbf{x}}^{i+1} = \arg\min_{\mathbf{x}}\left(\mathbf{y}-\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}\right)^*\Gamma_{\mathbf{y},\hat{\mathbf{x}}^i}^{-1}\left(\mathbf{y}-\sqrt{\rho}\mathbf{H}\mathbf{F}^*\mathbf{x}\right) \tag{4.13}$$

where max was replaced by min by removing the minus sign. The objective in (4.13) can be interpreted as resulting from the system given by (4.1) where

the noise vector $\mathbf{w}$ consists of Gaussian random variables each with a different variance given by $\gamma_j, \forall j$. In other words, this problem is similar to the problem in Section 4.5.1 with perfect noise state information (NSI) where the states are specified by $\Gamma_{\mathbf{y},\hat{\mathbf{x}}^i}$. Thus, taking the FFT will just lead to ICI as described in Section 4.5.1. The exact solution of (4.13) still requires an exponential search over $\mathbf{x}$. However; by formulating the problem as an EM problem, I transformed the highly non-linear objective of (4.10) into a quadratic objective given in (4.13). In addition, the problem was transformed from detection with no NSI (highly non-linear) into multiple iterations of detection with perfect NSI (with linear MMSE estimate). As a result, I approximate the solution of (4.13) by taking the MMSE estimate of the OFDM symbol in the time domain using the NSI followed by hard detection similar to the method given in Section 4.5.1. As a result, the new step is given by

$$\hat{\mathbf{x}}^{i+1} \approx \left[\mathbf{F}\hat{\mathbf{u}}^{\mathbf{i+1}}\right] \tag{4.14}$$

where $[\cdot]$ denotes hard detection and $\hat{u}^{i+1}$ is given by its Linear MMSE estimate as follows

$$\hat{u}^{i+1} = \text{diag}\left\{\frac{\|\mathbf{h}\|^2\rho}{\|\mathbf{h}\|^2\rho+\gamma_1^i}, \cdots, \frac{\|\mathbf{h}\|^2\rho}{\|\mathbf{h}\|^2\rho+\gamma_N^i}\right\}\mathbf{y}. \tag{4.15}$$

The choice of the initial value $\hat{\mathbf{x}}^0$ for the EM algorithm has a big effect on the convergence rate and converging value. Two possible initial points are: 1) the result of the typical OFDM receiver (taking an FFT followed by hard decision), and 2) taking the result of the MMSE receiver without NSI described in Section 4.5.2. The former is computationally more tractable since it involves

only an FFT operation while the latter might provide a better estimate and lead to lower number of iterations. This is explored further in the results section.

4.8 Numerical Results

The communication performance of the discussed algorithms is compared for $N = 1024$ with 4-QAM modulation in the presence of a 2-term Gaussian mixture model (also called the ϵ-contaminated Gaussian). In practice, the 2-term approximation is usually sufficient [25, 75]. For a fair comparison with single carrier systems, I assume I have a flat fading identity channel; *i.e.* $\mathbf{H} = \mathbf{I}$. The symbol error rate (SER) of the proposed iterative method is given for the conventional OFDM and single carrier (SC) receivers and for the non-iterative estimator-correlator receivers with NSI and without NSI. The noise parameters are set to the following typical values: $\pi = \begin{bmatrix} 0.9 & 0.1 \end{bmatrix}$, $\sigma_1^2 = 1$, and $\sigma_2^2 = 150$. SNR is defined as the signal power to the second moment of the impulsive noise.

4.8.1 Single Carrier vs. Conventional OFDM

Figure 4.1 shows the communication performance degradation between single carrier (SC) systems and conventional OFDM receivers (using FFT followed by hard detection). It is noticed that the single carrier system performs better at low SNR till around 6dB. After that the conventional OFDM system considerably outperforms the SC system with gains up to 7.5dB at SER=10^{-4}.

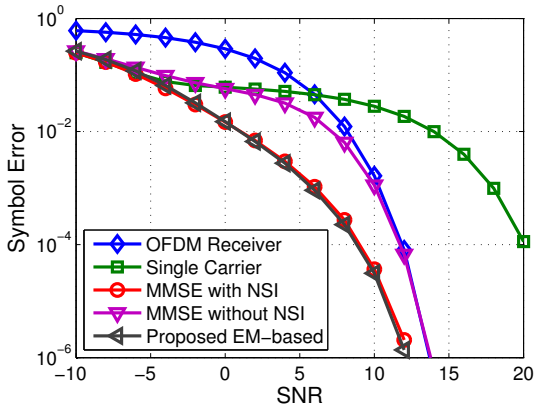


Figure 4.1: Communication performance of the low-complexity receivers in the presence of impulsive noise ($\pi_1 = 0.9$, $\pi_2 = 0.1$, $\sigma_1^2 = 1$ and $\sigma_2^2 = 100$). The proposed method has a gain of around 6dB in the moderate SNR region over the next best implementable algorithm.

This can be explained by the fact that at low SNR the occurring impulses have a much larger energy than the signal. In SC systems, this translates to losing the symbol exposed to the impulse. However, in OFDM systems the high energy of the impulse is spread across the whole OFDM symbol which results in losing the whole OFDM block. As a result, the SC performs better at low SNR. The opposite occurs at high SNR where the amplitude of the impulse is spread across the whole OFDM symbol without affecting it, while the SC system still suffers from the single symbols errors as in the previous case. This is the basis for the OFDM impulse resilience ability which is the result of the time diversity it provides when viewed as a time-code.

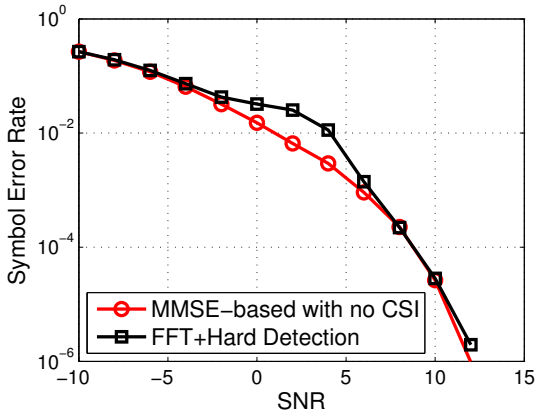


Figure 4.2: Communication performance for different initial values of $\mathbf{x}^0$ with 10 iterations of the EM algorithm. The initial value obtained by the MMSE-based detector with no NSI provides a slightly better performance for additional computational complexity than the conventional OFDM receiver.

4.8.2 Performance of the Proposed Method

The communication performance of the non-itarative MMSE methods described in Section 4.5 and the proposed iterative method based on the EM algorithm are shown in Figure 4.1. The non-iterative method with NSI provides the lower bound on the achievable performance using these time-domain MMSE based class of algorithms. However, the perfect NSI at the receiver assumption is not valid in most cases and this algorithm is impractical. The iterative algorithm is allowed to run for a maximum of 10 iterations. It is seen that the proposed EM-based algorithm provides a gain of ranging from 2dB to 7dB over the non-iterative MMSE without NSI. Further, the proposed method, which is an approximate ML detector, achieves the lower bound for the MMSE based methods with perfect NSI at almost the same computational

complexity. The effect of the choice of the initial condition on the performance of the proposed EM-based algorithm is given in Figure 4.2. The MMSE-based detector with no NSI provides only a slight improvement at 10 iterations of the EM algorithm.

4.9 Computational Complexity and Limitations

The computational complexity of the proposed algorithm is analyzed in terms of the number of exponential evaluations and FFT operations it requires per subcarrier. For each subcarrier k, the proposed algorithm has to compute γ_k given in (4.12) for each iteration. This requires K (number of Gaussian components, usually 2) scalar exponential evaluations that could be implemented in a lookup table. In addition to that, at the end of each iteration a FFT operation of size N has to be performed. Although the proposed method leads to significant performance gain for typical impulsive environments, it can fail in impulsive noise with extreme amplitudes that are for example 30dB higher above the noise floor. Such scenarios are not common since in many cases very high impulses are clipped by the receiver reducing their amplitude. This degradation in performance is due to the approximation made in (4.14) by which the EM algorithm loses its monotonic increase in likelihood. In addition, due to the independent subcarrier decoding, it doesn't take full advantage of OFDM's code diversity and its performance lags behind the theoretical lower bounds derived in Chapter 5 and [45, 46]. Another shortcoming is that this algorithm doesn't explicitly consider channel estimation; thus it will utilize

the suboptimal LMMSE channel estimator.

4.10 Conclusion

Using a latent variable interpretation of the interference model, I propose an EM-based OFDM receiver for impulsive noise channels that is constrained to perform independent subcarrier decoding. Compared to the conventional DFT receiver, the proposed receiver provides a gain of around 6dB in the low and moderate SNR range and about 2dB in the high SNR range. To achieve this gain, the EM receiver uses two aspects of the communication system: 1) OFDM modulated signals can be approximated as being iid Gaussian in time-domain (by the central limit theorem); and 2) the knowledge of the interference model pdf and its parameters. When combined under the EM framework, these facts lead to a simple scalar LMMSE estimator in the time-domain followed by DFT detection in frequency domain. While computationally attractive, such a disjoint decoding doesn't allow this receiver to fully exploit the OFDM's resilience to impulsive interference. Furthermore, as most of the prior work, this EM framework doesn't explicitly consider channel estimation. This can be a practical limitation as the conventional LMMSE channel estimation is highly suboptimal in non-AWGN environments.

To address the limitations of EM-based receiver, I propose in the following chapter a fully Bayesian inference framework to design LDPC-coded OFDM receivers in uncoordinated interference using the GM and GHMM models as priors. In particular, I propose a factor-graph-based approach to joint

channel/noise-estimation-and-decoding (JCNED) of orthogonal frequency division multiplexing (OFDM) systems in interference-limited environments. My receiver merges prior knowledge of the impulsive and bursty noise models with the recently proposed "generalized approximate message passing" (GAMP) algorithm, and soft-input soft-output decoding through the sum-product framework.

4.11 Appendix

The system model of (4.1) can be expressed as

$$y_j = u_j + w_j \qquad j = 1, \cdots, T.$$

The MMSE estimate of u_j given y_j is

$$\hat{u}_j = \mathbb{E}\left\{u_j | y_j\right\} = \int_{\mathbb{C}^N} u_j p\left(u_j | y_j\right) \mathrm{d}u_j. \tag{4.16}$$

Using Bayes rule and summing over all noise state realizations $s_j \in \mathcal{S}$ with $P(s_j) = \pi_j$, I obtain

$$\begin{aligned} p\left(u_j | y_j\right) &= \frac{\sum_{\mathcal{S}} \pi_j p\left(u_j | y_j, s_j\right) p\left(y_j | s_j\right)}{\sum_{\mathcal{S}} \pi_j p\left(y_j | s_j\right)} \\ &= \frac{\mathbb{E}_s\left\{p\left(u_j | y_j, s_j\right) p\left(y_j | s_j\right)\right\}}{\mathbb{E}_s\left\{p\left(y_j | s_j\right)\right\}} \end{aligned} \tag{4.17}$$

Substituting (4.17) in (4.16) and interchanging the order of integration and expectation

$$\hat{u}_j = \frac{\mathbb{E}_s\left\{p\left(y_j | s_j\right) \cdot \int u_j p\left(u_j | y_j, s_j\right) \mathrm{d}u_j\right\}}{\mathbb{E}_s\left\{p\left(y_j | s_j\right)\right\}}.$$

Given the noise state s_j, y_j is a sum of two independent Gaussian vectors and therefore Gaussian with covariance $\rho + \|\mathbf{h}\|^2\sigma_{s_j}^2$. On the other hand, $\int u_j p\left(u_j|y_j, s_j\right) \mathrm{d}u_j = \mathbb{E}\left\{u_j|y_j, s_j\right\}$ is the LMMSE estimate of u_j given in (4.7). Thus,

$$\hat{u}_j = \frac{\mathbb{E}_s\left\{\frac{1}{\|\mathbf{h}\|^2\rho+\sigma_s^2}\exp\left(-\frac{\|y_j\|^2}{\|\mathbf{h}\|^2\rho+\sigma_s^2}\right)\cdot\frac{\|\mathbf{h}\|^2\rho}{\|\mathbf{h}\|^2\rho+\sigma_{s_j}^2}y_j\right\}}{\mathbb{E}_s\left\{\frac{1}{\|\mathbf{h}\|^2\rho+\sigma_s^2}\exp\left(-\frac{\|y_j\|^2}{\|\mathbf{h}\|^2\rho+\sigma_s^2}\right)\right\}}$$

which simplifies to (4.8).

Chapter 5

Message-Passing OFDM Receivers for Impulsive Noise Channels

In Chapter 4, I proposed an EM-based receiver that utilizes the interference GM model to perform disjoint subcarrier detection. This receiver typically leads to a 2dB to 7dB improvement in communication performance over DFT OFDM receivers discussed in Section 2.1. However, by not considering the interference dependance across subcarriers, its performance falls short from the huge performance gains predicted by the PEP analysis done in [45, 46]. In this chapter, I propose a fully Bayesian inference framework to design OFDM receivers in uncoordinated interference by using the GM and GHMM models developed in Chapter 3 as priors. In particular, I propose a factor-graph-based approach to joint channel/noise-estimation-and-decoding (JCNED) of orthogonal frequency division multiplexing (OFDM) systems in interference-limited environments. My receiver merges prior knowledge of the impulsive and bursty noise models with the recently proposed "generalized approximate message passing" (GAMP) algorithm, and soft-input soft-output decoding through the sum-product framework. Unlike the prior work, I explicitly consider channel estimation in the problem formulation. For N subcarriers, the resulting receiver has a complexity of $\mathcal{O}(N \log N)$, comparable to

a typical DFT receiver, that can be parallelized and implemented efficiently on FPGAs. Numerical results indicate that the proposed receiver outperforms all prior impulsive noise OFDM decoders with improvements that reach 13dB when compared to the commonly used DFT receiver.

5.1 Introduction

As discussed in Section 2.1, the effect of OFDM modulation is to spread the impulsive noise energy across all frequency tones due to the discrete Fourier transform (DFT) applied at the receiver [36]. This effect is both a blessing and a curse: a blessing because it provides extra resilience by reducing the impulse energy experienced by each symbol, but a curse because the noise is no longer independent across sub-carriers and, unlike the AWGN case, independent decoding of each sub-channel is no longer optimal. In fact, under a wide range of operating conditions, the theoretical information limit of OFDM under independent sub-channel decoding is lower than for SC systems [78]. However, pairwise error probability (PEP) analysis of OFDM, under joint sub-channel decoding, demonstrates huge performance gains (up to 25 dB) over SC systems in impulsive noise channels [46, 47]. An additional challenge for wideband OFDM systems in impulsive noise, not considered in prior work or in Chapter 4, is the channel tap estimation; while the optimal estimator is linear under AWGN and independent Gaussian priors on $\{h_j\}_{j=0}^{L-1}$, this is not the case under impulsive noise or clustered-sparse channel taps which further complicates the receiver design [53, 85].

The prior work about OFDM receiver design in impulsive noise (discussed in Section 2.4) generally takes a *decoupled* and *suboptimal* approach to the problem of channel and impulse noise estimation and data decoding: first, the null tone knowledge is exploited to estimate the impulsive noise vector, after which it is subtracted from the received signal and passed to the DFT receiver which involves a linear estimation of the channel taps using pilot tones followed by data decoding.

5.2 Contribution

In this chapter, I show that such an approach is suboptimal by proposing a novel model-based low-complexity decoder *jointly* decodes the information bits and estimates the channel and the impulsive noise vector utilizing the impulsive noise models and the information available in all OFDM tones (not just null or pilot tones). This receiver achieves significant performance gains (tens of dBs) under practical impulsive noise channels. In particular, my receiver leverages recent results in "generalized approximate message passing" (GAMP) [79], soft-input/soft-output (SISO) decoding [59], and structured sparse estimation [80] to achieve an implementation complexity of only $\mathcal{O}\left(N \log N\right)$ where N denotes the number of sub-carriers of the OFDM system. The resulting implementation is order of magnitudes faster than competing receivers and can be parallelized providing a natural mapping to FPGA implementations.

5.3 System Model

5.3.1 Coded OFDM Model

I consider an N-tone OFDM system with the following tone partition: N_{p} pilot tones indexed by the set $\mathcal{P}$, N_{n} null tones indexed by $\mathcal{N}$, and N_{d} data tones indexed by the set $\mathcal{D}$ where each data subcarrier is modulated by a symbol from an 2^M-ary constellation $\mathbb{S}$. The data bits which are mapped to the data symbols are generated by encoding M_{i} information bits using a rate-R coder, interleaving them, and allocating the resulting $M_{\mathsf{c}} = M_{\mathsf{i}}/R$ bits among an integer number $Q = \lceil M_{\mathsf{c}}/N_{\mathsf{d}}M \rceil$ of OFDM symbols.

In the sequel, I use $s^{(i)} \in \mathbb{S}$ for $i \in \left\{1, \cdots, 2^M\right\}$ to denote the ith element of $\mathbb{S}$, and $\mathbf{c}^{(i)} = [c_1^{(i)}, \cdots, c_M^{(i)}]^T$ to denote the corresponding bits as defined by the symbol mapping. Likewise, I use $s_k[q]$ to denote the symbol transmitted on the kth subcarrier of the qth OFDM symbol. Based on the tone partition, I note that: $s_k[q] = \mathsf{p}$ for all $k \in \mathcal{P}$, where p is a known pilot symbol; $s_k[q] = 0$ for all $k \in \mathcal{N}$; and $s_k[q] = s^{(l)}$ for some l such that $\mathbf{c}_k[q] = \mathbf{c}^{(l)}$ for all $k \in \mathcal{D}$, where $\mathbf{c}_k[q] = [c_{k,1}[q], \cdots, c_{k,M}[q]]^T$ denotes the coded/interleaved bits corresponding to $s_k[q]$. On the frame level, I use $\mathbf{c}[q]$ to denote the coded/interleaved bits allocated to the data tones of the qth OFDM symbol, and $\mathbf{c} = [\mathbf{c}[1], \cdots, \mathbf{c}[Q]]^T$ to denote the entire codeword obtained from the information bits $\mathbf{b} = [b_1, \cdots, b_{M_{\mathsf{i}}}]^T$ by coding/interleaving.

OFDM modulation applies an N-point inverse discrete Fourier transform (IDFT) $\mathbf{F}^*$ to a vector of N symbols. These symbols are said to be in the frequency-domain since they are recovered by applying the DFT $\mathbf{F}$. The

resulting time-domain signal corresponding to the qth OFDM symbol is given by $\mathbf{u}[q] = \mathbf{F}^*\mathbf{s}[q]$ where $\mathbf{u}[q] = [u_0[q] \cdots u_{N-1}[q]]^T$ is the time domain signal and $\mathbf{s}[q] = [s_0[q] \cdots s_{N-1}[q]]^T$ is the transmitted symbol sequence.

After appending the cyclic prefix, the qth OFDM symbol's waveform propagates through a noisy LTI channel with a channel impulse response $\mathbf{h}[q] = [h_0[q] \cdots h_{L-1}[q]]^T$ where L is the number of channel taps. After discarding the cyclic prefix and assuming perfect synchronization at the receiver, the received signal $\mathbf{r}[q]$ can be expressed as

$$\mathbf{r}[q] = \mathbf{H}[q]\mathbf{u}[q] + \mathbf{n}[q] = \mathbf{H}[q]\mathbf{F}^*\mathbf{s}[q] + \mathbf{n}[q] \tag{5.1}$$

where $\mathbf{n}[q]$ is additive noise, and $\mathbf{H}[q]$ is the circulant matrix formed by $\mathbf{h}[q]$ [94]. The receiver applies DFT to $\mathbf{r}[q]$, and the resulting frequency-domain signal is given by

$$\mathbf{y}[q] = \mathbf{F}\mathbf{r}[q] = \mathbf{F}\mathbf{H}[q]\mathbf{F}^*\mathbf{s}[q] + \mathbf{F}\mathbf{n}[q] = \mathsf{H}[q] \circ \mathbf{s}[q] + \mathsf{N}[q] \tag{5.2}$$

where $\mathsf{H}[q] = (\sqrt{N}\mathbf{F}_{:,1:L})\mathbf{h}[q]$ denotes the frequency-domain channel, $\mathsf{N}[q] = \mathbf{F}\mathbf{n}[q]$ the frequency-domain noise, and $\circ$ denotes the Hadamard product. In short, (5.2) illustrates the main advantage of OFDM modulation: each tone k now experiences a flat fading channel given by

$$y_k[q] = \mathsf{H}_k[q]s_k[q] + \mathsf{N}_k[q], \quad \forall k \in \{0, \cdots, N-1\}. \tag{5.3}$$

To simplify the development, I assume that $Q = 1$ in the sequel (but not in the simulations), and drop the index $[q]$ for brevity.

5.3.2 Impulsive Noise Models

Since my message-passing receiver is inherently Bayesian, the statistical-physical models presented in Chapter 3 provide natural priors on the impulsive noise.I note that, given the pdf parameters, there is no distinction between the MCA and GM models from the receiver design perspective. I now review the models used in this chapter.

The pdf of a GM distributed random variable n with zero mean components is given by

$$p\left(n\right)=\sum_{k=0}^{K-1}\pi^{(k)}\cdot\mathcal{N}\left(n;0,\gamma_k\right) \tag{5.4}$$

where $\mathcal{N}\left(n;0,\gamma_k\right)$ denotes the complex Gaussian pdf with zero mean and variance γ_k; and $\pi^{(k)}$ is the probability of the k-th Gaussian component. Generally, the Gaussian background noise $g\sim\mathcal{N}\left(0,\gamma_0\right)$ is the component with the smallest variance, and the total noise can be decomposed as the sum of background and impulsive noise as $n=g+i$, where i is the impulsive noise component whose pdf is given by

$$p\left(i\right)=\pi^{(0)}\delta\left(i\right)+\sum_{k=1}^{K-1}\pi^{(k)}\cdot\mathcal{N}\left(i;0,\gamma_k\right). \tag{5.5}$$

Eq. (5.5) admits a simple interpretation: with probability $\pi^{(0)}$ only background noise is present and with probability $1-\pi^{(0)}$ I have a non-zero impulsive noise on top of the background noise. Here, the variances γ_k indicate the power of the impulse component rather than the total variance of the noise n as in (5.4) (γ_k in (5.4) equals $\gamma_0+\gamma_k$ in (5.5)). The mixing vector $\boldsymbol{\pi}=[\pi^{(0)}\cdots\pi^{(K-1)}]$

can be interpreted as a probability mass function (pmf) of a latent random variable z that selects the mixture component used to generate a sample of n; *i.e.* $P(z = k) = \pi^{(j)}, \forall k \in \{0, \cdots, K-1\}$. Furthermore, given z, I have $i|\{z = k\} \sim \mathcal{N}(0, \gamma_k)$ where $\gamma_0 = 0$ corresponds to the absence of impulsive noise. On the other hand, the Gaussian HMM (GHMM) captures the temporal dynamics of bursty noise by embedding the latent state variable z_t into a Markov chain with a state transition matrix $\mathbf{T}$ defined as

$$\mathbf{T}_{i,j} = P(z_t = j|z_{t-1} = i) \quad \forall i, j \in \{0, \cdots, K-1\} \tag{5.6}$$

where I assumed that the Markov chain is homogeneous (stationary). I note that the marginal distribution of the noise samples under this model is the GM distribution given in (5.5) with $\boldsymbol{\pi}$ given by the solution of $\mathbf{x} = \mathbf{x}\mathbf{T}$. Similar to the GM model, conditioning on the state $z_t = k$ makes the noise samples independent and Gaussian with variance γ_k. Under this model, the mean duration of a bursty impulse noise from state k is given by $\mathbb{E}\{\tau_k\} = 1/(1 - \mathbf{T}_{k,k})$ where τ_k is the duration that the chain will persist in state k [32].

5.4 Decoding in Impulsive Noise Channels

My objective is to infer the information bits $\mathbf{b}$ transmitted on the data tones given the received signal $\mathbf{y}$, the prior knowledge about the impulsive noise and the channel, and the locations of the pilot and null tones given by $\mathcal{P}$ and $\mathcal{N}$, respectively. A fully Bayesian approach first marginalizes over the

channel taps $\mathbf{h}$ and the noise vector $\mathbf{n}$, considered latent variables, and then produces the likelihood ratio of each information bit b_m. The MAP detection rule can then be formulated as

$$\hat{b}_m = \underset{b_m \in \{0,1\}}{\arg\max}\, P(b_m|\mathbf{y}, \Theta)\,, \quad \forall m \in \{1, \cdots, M_{\mathsf{i}}\} \tag{5.7}$$

where $\Theta = \{\mathcal{D}, \mathcal{P}, \mathcal{N}, \theta_{\mathbf{i}}, \theta_{\mathbf{h}}\}$, $\theta_{\mathbf{i}} = \{\mathbf{T}, \{\gamma_k\}_{k=0}^{K-1}\}$ is the noise model parameters, and $\theta_{\mathbf{h}}$ is the channel model parameters.

To highlight the challenge of data decoding in impulsive noise, recall the OFDM system model given by

$$\mathbf{y} = \mathsf{H} \circ \mathbf{s} + \mathbf{F}\mathbf{g} + \mathbf{F}\mathbf{i} = \mathsf{H} \circ \mathbf{s} + \mathsf{I} + \mathsf{G} \tag{5.8}$$

where I have explicitly written the noise as a sum of background and impulsive noise. While, G is also an AWGN vector with independent components, this is not the case for I. To see this, consider taking the DFT of a zero vector with a single impulse; the components of the resulting vector are all non-zero with amplitude that is a function of the impulse and its location. Due to this coupling, the independent and disjoint decoding of the OFDM subchannels as done under AWGN is no longer optimal.

Using the *law of total probability*, the posterior info-bit probability used

in (5.7) can be written as

$$
\begin{aligned}
P(b_m|\mathbf{y};\Theta) &= \sum_{\mathbf{b}_{\backslash\mathbf{m}}} P(\mathbf{b}|\mathbf{y};\Theta) \propto \sum_{\mathbf{b}_{\backslash\mathbf{m}}} P(\mathbf{y}|\mathbf{b};\Theta)\, P(\mathbf{b}) \\
&= \sum_{\mathbf{s},\mathbf{c},\mathbf{b}_{\backslash\mathbf{m}}} \int_{\mathbf{i},\mathbf{h}} p\left(\mathbf{y}|\mathbf{h},\mathbf{i},\mathbf{s};\Theta\right) p\left(\mathbf{h};\theta_{\mathbf{h}}\right) p\left(\mathbf{i};\theta_{\mathbf{i}}\right) \\
&\quad \times P(\mathbf{s}|\mathbf{c})\, P(\mathbf{c}|\mathbf{b})\, P(\mathbf{b}) \\
&= \sum_{\mathbf{s},\mathbf{c},\mathbf{z},\mathbf{d},\mathbf{b}_{\backslash\mathbf{m}}} \int_{\mathbf{i},\mathbf{h}} \prod_{k=0}^{N-1} p\left(y_k|s_k,\mathbf{h},\mathbf{i};\mathcal{D},\mathcal{N},\mathcal{P}\right) \\
&\quad \times P(s_k|\mathbf{c}_k)\, p\left(i_k|z_k\right) P(z_k|z_{k-1})\, P(\mathbf{c}|\mathbf{b}) \\
&\quad \times \prod_{l=0}^{L-1} p\left(h_l|d_l\right) P(d_l|d_{l-1}) \prod_{i=1}^{M_{\mathrm{i}}} P(b_i)
\end{aligned} \tag{5.9}
$$

where $\mathbf{b}_{\backslash\mathbf{m}} = [b_1, \cdots, b_{m-1}, b_{m+1}, \cdots, b_{M_{\mathrm{i}}}]^T$ and d_l is the latent tap state for clustered-tap channels (see [85]). The coupling between the subchannels is evident in the $p\left(y_j|s_j,\mathbf{h},\mathbf{i},\mathcal{D},\mathcal{P},\mathcal{N}\right)$ term where the received signal at each tone y_j depends on the complete vectors $\mathbf{i}$ and $\mathbf{h}$ through the linearly mixed terms $\mathsf{I}_j = [\mathbf{Fi}]_j$ and $\mathsf{H}_j = [\mathbf{Fh}]_j$, respectively. This prevents the high-dimensional integrals in (6.19) from simplifying into N scalar integrals, as would happen under AWGN, and leads to an intractable marginalization.

5.5 Message Passing Receivers

The factorization of the pdf given in (6.19) is represented by the *factor graph* in Figure 5.1 where the round nodes denote random variables and the square nodes denote the factors of the posterior. The factor node y_k, representing $p\left(y_k|\mathsf{H}_k,\mathsf{N}_k,s_k,\Theta\right)$, connects to the transmitted symbol s_k (possibly a

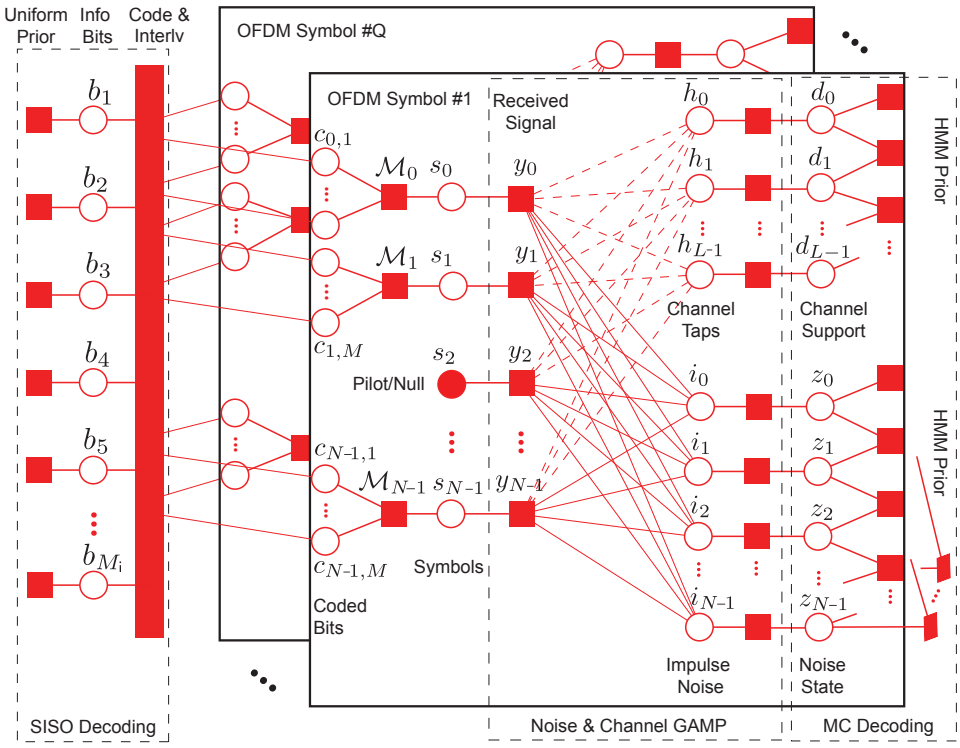


Figure 5.1: Factor graph representation of a coded data frame allocated across Q OFDM symbols: the dense sub-graphs formed between the factors $\mathbf{y}$ and time-domain $\mathbf{i}$ and $\mathbf{h}$ are due to the linear mixing via the Fourier matrix $\mathbf{F}$; solid circles (s_3) represent known tones (pilots or nulls).

null or pilot tone) and, as a result of the linear mixing via $\mathbf{F}$, to every impulsive noise sample and channel tap in the time-domain forming two dense sub-graphs: one with the impulsive noise samples and another with the channel taps.

I now provide the background on factor graph inference followed by the derivation of my message-passing receivers.

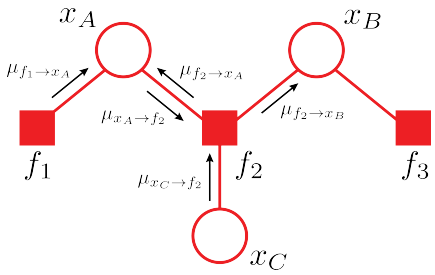


Figure 5.2: The messages $\mu_{\cdot \to \cdot}$ propagate beliefs used for inference. By the sum-product algorithm, $\mu_{x_A \to f_2} = \mu_{f_1 \to x_A}$ and $\mu_{f_2 \to x_B} = \sum_{x_A, x_C} p\left(x_C | x_A, x_B\right) \mu_{x_A \to f_2}\left(x_A\right) \mu_{x_C \to f_2}\left(x_C\right)$.

5.5.1 Belief Propagation using Sum-Product Algorithm

Inference using belief propagation (BP) transforms a high-dimensional marginalization problem into a series of local low-dimensional marginalization by passing beliefs in the form of pdf/pmf messages along the edges of a factor graph using the sum-product (SP) algorithm [12]. Let us consider the factor graph of Section 2.5.2 corresponding to the following factorization of a given pdf

$$p\left(x_A, x_B, x_C\right) = p\left(x_A\right) p\left(x_B\right) p\left(x_C | x_A, x_B\right) \tag{5.10}$$

and shown again in Figure 5.2. The SP algorithm computes the messages as follows:

5.5.1.1 Messages from Factor Nodes to Variables

The message passed from a factor node f_s to a variable x_m is given by

$$\mu_{f_s \to x_m}\left(x_m\right) = \int_{\mathbf{x_s}} f_s\left(x_m, \mathbf{x_s}\right) \prod_{x_l \in \mathrm{ne}(f_s) \setminus x_m} \mu_{x_l \to f_s}\left(x_l\right) \, \mathrm{d}\mathbf{x}_s$$

where $\mathbf{x_s} = [\mathrm{ne}(f_s) \setminus x_m]$ is a vector of variable nodes that are neighbors of f_s excluding the recipient variable x_m. This message represents the belief of the factor node f_s about the variable x_m that is obtained by marginalizing the pdf factor represented by f_s over all variables in $\mathrm{ne}(f_s) \setminus x_m$ using the beliefs passed on to f_s about those variables.

5.5.1.2 Messages from Variables to Factor Nodes

The message passed from a variable node x_m to a factor node f_s is

$$\mu_{x_m \to f_s}(x_m) = \prod_{f_k \in \mathrm{ne}(x_m) \setminus f_s} \mu_{f_k \to x_m}(x_m).$$

This can be interpreted as passing an independent combination of the beliefs that factors in $\mathrm{ne}(x_m) \setminus f_s$ have about the variable x_m to the factor f_s.

5.5.1.3 Marginal Approximation

The marginal at any variable node x_m can be approximated as

$$p(x_m) = C \prod_{f_s \in \mathrm{ne}(x_m)} \mu_{f_s \to x_m}(x_m)$$

where C is a normalization constant. In other words, the beliefs from each factor node connected to x_m are treated as being independent; thus, their joint is the product of the individual messages.

When the factor graph doesn't contain any loops, BP-SP algorithm performs exact inference after only two rounds of messages (*i.e.* forward and backward passes). On the other hand, with loops in the factor graph, neither

convergence nor exact inference are guaranteed; nevertheless, this loopy BP has been successfully applied in many important problems: multi-user detection [16, 43], turbo decoding [63], LDPC decoding [59], and compressed sensing [10, 28, 79].

5.5.2 Approximate Message Passing

As discussed in Section 5.4, the main challenge in computing the symbol posteriors is the high dimensional inference problem due to the coupling introduced by the linear mixing between the frequency-domain symbols and time-domain noise and channel vectors via the Fourier matrix $\mathbf{F}$. An important sub-problem in this inference is the estimation of a vector of independent possibly-non-Gaussian variables $\mathbf{x}$ that are linearly mixed via $\Phi \in \mathcal{C}^{M \times N}$ to form $\mathbf{z} = \Phi\mathbf{x} = [z_1 \cdots z_M]^T$ and subsequently observed at the output of independent and possibly non-Gaussian channels with pdfs $\{p(y_i|z_i)\}_{i=1}^{M}$. The inference under such a model was addressed by the *generalized approximate message passing* (GAMP) algorithm proposed by Rangan [79, 80] as a generalization of relaxed BP by Guo and Wang [43] and approximate message passing (AMP) algorithm for compressed sensing by Donoho, Maleki, Montanari, and Bayati [11, 28].

The GAMP algorithm exploits the large dimensionality of the problem and the central limit theorem to approximate both the messages flowing leftward from the nodes $\{x_j\}_{j=1}^{N}$ and rightward from the factors $\{p(y_i|z_i)\}_{i=1}^{M}$ as Gaussian. The GAMP$(\mathbf{x}, \mathbf{z}, \Phi)$ algorithm, where $\mathbf{z} = \Phi\mathbf{x}$, is given in Ta-

ble 5.1. The detailed derivation and theoretical guarantees of the GAMP algorithm are beyond the scope of this dissertation; I refer the interested reader to [79] and [11] for more information. Here, I focus on the GAMP aspects that are specific to the design of my proposed receivers. In particular, the product of messages coming into a factor node $f_i = p\left(y_i|z_i\right)$ is approximated as $\prod_j \mu_{x_j \to f_i} \approx \mathcal{N}\left(z_i; \hat{p}_i, \gamma_i^p\right)$ where $\hat{p}_i$ and γ_i^p are given in (R3) and (R2) of Table 5.1. To compute the message from this factor node f_i to a node x_j, the approximated product is multiplied by the factor $p\left(y_i|z_i\right)$ according to SP algorithm rules in Section 5.5.1.1; however, instead of marginalizing over the remaining $\mathbf{x}\backslash x_j$, GAMP approximates the outgoing message using a second order Taylor series expansion summarized by two parameters: a mean $\hat{s}_i$ and variance γ_i^s given in (R4) and (R5) of Table 5.1. The product of these messages arriving at node x_j is further approximated as $\prod_i \mu_{f_i \to x_j} \approx \mathcal{N}\left(x_j; \hat{r}, \gamma_j^r\right)$ where the mean $\hat{r}$ and variance γ_j^r are computed from parameters $\hat{s}_i$ and γ_i^s given in (R6) and (R7) of Table 5.1. Next, this approximated message is multiplied by the prior factor $p\left(x_j\right)$ according to the SP update rule given in Section 5.5.1.2 and approximated as Gaussian using second order Taylor series expansion leading to a estimate of the posterior $p\left(x_j|y\right) = \mathcal{N}\left(x_j; \hat{x}_j, \gamma_j^x\right)$ where $\hat{x}_j$ and γ_j^x are given in (R8) and (R9) of Table 5.1. These parameters are then used in the computations of $\hat{p}$ and γ^p in (R2-3) thereby completing a single GAMP iteration. From (D2-3) and (D5-6) in Table 5.1, the GAMP algorithm requires the derivation of four scalar functions that are a function of the effective observation channels and signal priors: $g_{\mathsf{out},i}(y_i, \hat{z}, \gamma^z)$, $g_{\mathsf{in},j}(\hat{r}, \gamma^r)$,

and their derivatives with respect of $\hat{z}$ and $\hat{r}$, respectively.

Table 5.1: The GAMP$(\mathbf{x}, \mathbf{z}, \Phi)$ Algorithm

Definitions:			
$p\left(z_i \vert y_i; \hat{z}, \gamma^z\right)$	$=$	$\frac{p(y_i \vert z_i)\,\mathcal{N}(z_i;\hat{z},\gamma^z)}{\int_{z_i} p(y_i \vert z_i)\,\mathcal{N}(z_i;\hat{z},\gamma^z)}$	(D1)
$g_{\mathsf{out},i}(y_i, \hat{z}, \gamma^z)$	$=$	$\frac{1}{\gamma^z}\left(\mathbb{E}\left\{z_i \vert y_i; \hat{z}, \gamma^z\right\} - \hat{z}\right)$	(D2)
$g'_{\mathsf{out},i}(y_i, \hat{z}, \gamma^z)$	$=$	$\frac{1}{\gamma^z}\left(\frac{\mathbb{V}\{z_i \vert y_i;\hat{z},\gamma^z\}}{\gamma^z} - 1\right)$	(D3)
$p\left(x_j \vert \hat{r}; \gamma^r\right)$	$=$	$\frac{p(x_j)\,\mathcal{N}(x;\hat{r},\gamma^r)}{\int_{x_j} p(x_j)\,\mathcal{N}(x_j;\hat{r},\gamma^r)}$	(D4)
$g_{\mathsf{in},j}(\hat{r}, \gamma^r)$	$=$	$\mathbb{E}\left\{x_j \vert \hat{r}; \gamma^r\right\}$	(D5)
$g'_{\mathsf{in},j}(\hat{r}, \gamma^r)$	$=$	$\frac{1}{\gamma^r}\mathbb{V}\left\{x_j \vert \hat{r}; \gamma^r\right\}$	(D6)
Initialize:			
$\forall j : \hat{x}_j(1)$	$=$	$\int_x x\, p_{X_j}(x)$	(I1)
$\forall j : \gamma_j^x(1)$	$=$	$\int_x \lvert x - \hat{x}_j(1)\rvert^2 p_{X_j}(x)$	(I2)
$\forall i : \hat{u}_i(0)$	$=$	0	(I3)
for $t = 1, 2, 3, \ldots$			
$\forall i : \hat{z}_i(t)$	$=$	$\sum_{j=1}^N \Phi_{ij}\hat{x}_j(t)$	(R1)
$\forall i : \gamma_i^p(t)$	$=$	$\sum_{j=1}^N \lvert\Phi_{ij}\rvert^2\gamma_j^x(t)$	(R2)
$\forall i : \hat{p}_i(t)$	$=$	$\hat{z}_i(t) - \gamma_i^z(t)\,\hat{u}_i(t-1)$	(R3)
$\forall i : \hat{s}_i(t)$	$=$	$g_{\mathsf{out},i}(y_i, \hat{p}_i(t), \gamma_i^z(t))$	(R4)
$\forall i : \gamma_i^s(t)$	$=$	$-g'_{\mathsf{out},i}(y_i, \hat{p}_i(t), \gamma_i^z(t))$	(R5)
$\forall j : \gamma_j^r(t)$	$=$	$\left(\sum_{i=1}^M \lvert\Phi_{ij}\rvert^2\gamma_i^s(t)\right)^{-1}$	(R6)
$\forall j : \hat{r}_j(t)$	$=$	$\hat{x}_j(t) + \gamma_j^r(t)\sum_{i=1}^M \Phi_{ij}^*\hat{s}_i(t)$	(R7)
$\forall j : \gamma_j^x(t+1)$	$=$	$\gamma_j^r(t)g'_{\mathsf{in},j}(\hat{r}_j(t), \gamma_j^r(t))$	(R8)
$\forall j : \hat{x}_j(t+1)$	$=$	$g_{\mathsf{in},j}(\hat{r}_j(t), \gamma_j^r(t))$	(R9)
end			

5.5.3 Joint Channel/Noise Estimation and Decoding (JCNED)

In designing my receiver, I take a fully Bayesian viewpoint and marginalize over the channel taps and the impulsive noise samples by performing loopy BP on the factor graph given in Figure 5.1. In particular, I make use of

the GAMP algorithm to tackle the high dimension inference presented by the impulsive noise and the channel dense sub-graphs.

Given the loopy nature of the factor graph, there exists considerable freedom in the message passing schedule. For JCNED, I choose to pass the messages roughly from right to left and backward in the following fashion: (i) starting at the info-bits I send messages toward the coded bits and then to the symbols; (ii) I then pass the messages to the channel sub-graph which iterates *between* GAMP iterations and Markov chain (MC) channel-tap-state decoding[1] in what I refer to as "equalizer" iterations; (iii) after convergence, the computed messages are passed back to the $\mathbf{y}$ factor nodes which in turn pass them to the impulsive noise sub-graph so it can perform its own equalizer iterations and pass the result back to the factor nodes $\mathbf{y}$; (iv) finally, the messages are propagated back toward the coding/interleaving factor which performs SISO decoding and updates the info-bit likelihoods. I will refer to each of these full cycles as a "turbo" iteration. I note that it is also possible to execute a parallel schedule if the hardware platform supports it. The message passing details are discussed below.

Without any prior knowledge about the information bits, the info-bit belief flowing into the coding/interleaving factor node is $P(b_m) = 1/2, \forall m$. As a result, the coded-bit beliefs flowing rightward out of the coding/interleaving factor node and into the symbol mapping nodes are uniformly distributed. The

[1]Recall that MC decoding requires only one forward-backward iteration for loopy BP to converge since it doesn't contain any loops.

Table 5.2: The GAMP output scalar estimation functions used for channel inference in JCNED

Tone Type	**Channel Output Scalar Estimation Functions**
	$g_{\mathsf{out},k}(y_k, \hat{\mathsf{H}}, \gamma^{\mathsf{H}})$
Pilot: $k \in \mathcal{P}$	$\mathsf{p}^*(y_k - \hat{\mathsf{I}}_k - \mathsf{p}\hat{\mathsf{H}})/(\gamma_0 + \gamma_k^{\mathsf{I}} + \rho_{\mathsf{p}}\gamma^{\mathsf{H}})$
Data: $k \in \mathcal{D}$	$\sum_{l=1}^{\lvert\mathbb{S}\rvert} \lambda_k^{(l)} s^{*(l)}(y_k - \hat{\mathsf{I}}_k - \hat{\mathsf{H}}s^{(l)})/(\gamma_0 + \gamma_k^{\mathsf{I}} + \lVert s^{(l)}\rVert^2\gamma^{\mathsf{H}})$ where $\lambda_k^{(l)} = p\left(y_k \lvert s^{(l)}\right)\beta_k^{(l)} / \sum_j p\left(y_k \lvert s^{(j)}\right)\beta_k^{(j)}$ and $p\left(y_k \lvert s^{(l)}\right) = \mathcal{N}\left(y_k; \hat{\mathsf{I}}_k + \hat{\mathsf{H}}s^{(l)}, \gamma_0 + \gamma_k^{\mathsf{I}} + \lVert s^{(l)}\rVert^2\gamma^{\mathsf{H}}\right)$
	$-g'_{\mathsf{out},k}(y_k, \hat{\mathsf{H}}, \gamma^{\mathsf{H}})$
Pilot: $k \in \mathcal{P}$	$\rho_{\mathsf{p}}/(\gamma_0 + \gamma_k^{\mathsf{I}} + \rho_{\mathsf{p}}\gamma^{\mathsf{H}})$
Data: $k \in \mathcal{D}$	$\lVert g_{\mathsf{out},k}(y_k, \hat{\mathsf{H}}, \gamma^{\mathsf{H}}) + \hat{\mathsf{H}}/\gamma^{\mathsf{H}}\rVert^2 +$ $\sum_{l=1}^{\lvert\mathbb{S}\rvert} \lambda_k^{(l)} \left[\lVert s^{(l)}\rVert^2/(\gamma_0 + \gamma_k^{\mathsf{I}} + \lVert s^{(l)}\rVert^2\gamma^{\mathsf{H}}) - \lVert\hat{\mathsf{H}}/\gamma^{\mathsf{H}}\right.$ $\left. + s^{*(l)}(y_k - \hat{\mathsf{I}}_k - \hat{\mathsf{H}}s^{(l)})/(\gamma_0 + \gamma_k^{\mathsf{I}} + \lVert s^{(l)}\rVert^2\gamma^{\mathsf{H}})\rVert^2\right]$

symbol mapping is deterministic and forms a constraint factor node given by $P\left(s^{(i)}|\mathbf{c}^{(j)}\right) = \delta_{i-j}$. According to the SP rules, the message passed rightward from the symbol mapping factor node "$\mathcal{M}_k$" is given by

$$\begin{aligned}\mu_{\mathcal{M}_k \to s_k}(s^{(i)}) &\propto \sum_{\mathbf{c}_k \in \{0,1\}^M} P\left(s^{(i)}|\mathbf{c}_k\right) \prod_{m=1}^{M} \mu_{c_{k,m} \to \mathcal{M}_k}(c_m) \\ &= \prod_{m=1}^{M} \mu_{c_{k,m} \to \mathcal{M}_k}(c_m^{(k)}) \end{aligned} \tag{5.11}$$

which in turn is copied forward as the message passed rightward from node s_k; i.e. $\mu_{\mathcal{M}_k \to s_k}(s^{(i)}) = \mu_{s_k \to y_k}(s^{(i)})$.

The next step in my message-passing schedule employs GAMP$(\mathbf{h}, \mathsf{H}, \sqrt{N}\mathbf{F})$ to approximate the messages going in and out of the channel dense subgraph. From Figure 5.1, I note two types of messages flowing into each factor

Table 5.3: The GAMP output scalar estimation functions used for impulse noise inference in JCNED

Tone Type	**Impulsive Noise Output Scalar Estimation Functions**
	$g_{\mathsf{out},k}(y_k, \hat{\mathsf{I}}, \gamma^{\mathsf{I}})$
Null: $k \in \mathcal{N}$	$(y_k - \hat{\mathsf{I}})/(\gamma_0 + \gamma^{\mathsf{I}})$
Pilot: $k \in \mathcal{P}$	$(y_k - \hat{\mathsf{I}} - \hat{\mathsf{H}}_k \mathsf{p})/(\gamma_0 + \gamma^{\mathsf{I}} + \rho_{\mathsf{p}} \gamma_k^{\mathsf{H}})$
Data: $k \in \mathcal{D}$	$\sum_{l=1}^{\lvert\mathbb{S}\rvert} \lambda_k^{(l)} (y_k - \hat{\mathsf{I}} - \hat{\mathsf{H}}_k s^{(l)})/(\gamma_0 + \gamma^{\mathsf{I}} + \lVert s^{(l)} \rVert^2 \gamma_k^{\mathsf{H}})$ where $\lambda_k^{(l)} = p\left(y_k \lvert s^{(l)}\right) \beta_k^{(l)} / \sum_j p\left(y_k \lvert s^{(j)}\right) \beta_k^{(j)}$ and $p\left(y_k \lvert s^{(l)}\right) = \mathcal{N}\left(y_k; \hat{\mathsf{I}} + \hat{\mathsf{H}}_k s^{(l)}, \gamma_0 + \gamma^{\mathsf{I}} + \lVert s^{(l)} \rVert^2 \gamma_k^{\mathsf{H}}\right)$
	$-g'_{\mathsf{out},k}(y_k, \hat{\mathsf{I}}, \gamma^{\mathsf{I}})$
Null: $k \in \mathcal{N}$	$1/(\gamma_0 + \gamma^{\mathsf{I}})$
Pilot: $k \in \mathcal{P}$	$1/(\gamma_0 + \gamma^{\mathsf{I}} + \rho_{\mathsf{p}} \gamma_k^{\mathsf{H}})$
Data: $k \in \mathcal{D}$	$\lVert g_{\mathsf{out},k}(y_k, \hat{\mathsf{I}}, \gamma^{\mathsf{I}}) + \hat{\mathsf{I}}/\gamma^{\mathsf{I}} \rVert^2 +$ $\sum_{l=1}^{\lvert\mathbb{S}\rvert} \lambda_k^{(l)} \left[1/(\gamma_0 + \gamma^{\mathsf{I}} + \lVert s^{(l)} \rVert^2 \gamma_k^{\mathsf{H}}) - \lVert \hat{\mathsf{I}}/\gamma^{\mathsf{I}} + \right.$ $\left. (y_k - \hat{\mathsf{I}} - \hat{\mathsf{H}}_k s^{(l)})/(\gamma_0 + \gamma^{\mathsf{I}} + \lVert s^{(l)} \rVert^2 \gamma_k^{\mathsf{H}}) \rVert^2 \right]$

node y_k that determine the GAMP output channel $p\left(y_k|\mathsf{H}_k\right)$: (i) the symbol beliefs from the symbol node s_k denoted by $\boldsymbol{\beta}_k = [\beta_k^{(1)}, \cdots, \beta_k^{(|\mathbb{S}|)}]$ where $\beta_k^{(i)} = \mu_{s_k \to y_k}(s^{(i)})$; and (ii) N messages flowing leftward from impulsive noise nodes $\mathbf{i}$ whose product is approximated as being $\mathcal{N}\left(\mathsf{I}_k; \hat{\mathsf{I}}_k, \gamma_k^{\mathsf{I}}\right)$ and computed by GAMP$(\mathbf{i}, \mathsf{I}, \mathbf{F})$ in the previous turbo iteration (in the first turbo iteration $\hat{\mathsf{I}}_k = 0$ and $\gamma_k^{\mathsf{I}} = \gamma^i$). From (5.3) and (5.8), the resulting GAMP$(\mathbf{h}, \mathsf{H}, \sqrt{N}\mathbf{F})$ output channels are given by

$$p\left(y_k|\mathsf{H}_k\right) = \begin{cases} \mathcal{N}\left(y_k; \mathsf{p}\mathsf{H}_k + \hat{\mathsf{I}}_k, \gamma_k^{\mathsf{I}} + \gamma_0\right) & \text{if } k \in \mathcal{P} \\ \sum_{l=1}^{|\mathbb{S}|} \beta_k^{(l)} \mathcal{N}\left(y_k; s^{(l)}\mathsf{H}_k + \hat{\mathsf{I}}_k, \gamma_k^{\mathsf{I}} + \gamma_0\right) & \text{if } k \in \mathcal{D} \end{cases}$$

with the corresponding output scalar estimation functions given in Table 5.2

with their corresponding derivations highlighted in Appendix 5.8.1 and Appendix 5.8.2. The input scalar estimation function $g_{\mathsf{in},j}(\hat{r},\gamma^r)$ corresponding to the clustered-tap channel prior is derived in [85]. Using these scalar estimation functions, the GAMP$(\mathbf{h},\mathsf{H},\sqrt{N}\mathbf{F})$ in Table 5.1 is iterated until it converges generating close approximations to the conditional means $\hat{h}$ and variances $\boldsymbol{\gamma}^h=[\gamma_0^h,\cdots,\gamma_{L-1}^h]^T$ given the observations $\mathbf{y}$ and the prior information. After that, the generated channel tap-states are passed rightward to the channel tap sub-graph for MC decoding. This is repeated for several channel equalizer iterations as detailed in [85].

Upon the termination of the equalizer iterations in the channel sub-graph, L messages are passed leftward from the channel taps node $\mathbf{h}$ toward each factor node y_k. The product of these messages is approximated by GAMP$(\mathbf{h},\mathsf{H},\sqrt{N}\mathbf{F})$ as $\mathcal{N}\left(\mathsf{H}_k;\hat{\mathsf{H}}_k,\gamma_k^{\mathsf{H}}\right)$. These messages in addition to the symbol beliefs $\boldsymbol{\beta}$ now determine the GAMP$(\mathbf{i},\mathsf{I},\mathbf{F})$ output channels which based on (5.3) and (5.8) can be written as

$$p\left(y_k|\mathsf{I}_k\right)=\begin{cases}\mathcal{N}\left(y_k;\mathsf{I}_k,\gamma_0\right) & \text{if } k\in\mathcal{N}\\ \mathcal{N}\left(y_k;\mathsf{p}\hat{\mathsf{H}}_k+\mathsf{I}_k,\rho_{\mathsf{p}}\gamma_k^{\mathsf{H}}+\gamma_0\right) & \text{if } k\in\mathcal{P}\\ \sum_{l=1}^{|\mathbb{S}|}\beta_k^{(l)}\mathcal{N}\left(y_k;s^{(l)}\hat{\mathsf{H}}_k+\mathsf{I}_k,\rho_{\mathsf{p}}\gamma_k^{\mathsf{H}}+\gamma_0\right) & \text{if } k\in\mathcal{D}\end{cases}$$

with the corresponding output scalar estimation functions given in Table 5.3. The input scalar estimation function $g_{\mathsf{in},j}(\hat{r},\gamma^r)$ for GAMP$(\mathbf{i},\mathsf{I},\mathbf{F})$ depends on the current impulsive noise belief which from Figure 5.1 can be expressed as

$$p\left(i_j\right)=\pi_j^{(0)}\delta\left(i_j\right)+\sum_{k=1}^{K-1}\pi_j^{(k)}\cdot\mathcal{N}\left(i_j;0,\gamma_k\right) \tag{5.12}$$

where $\boldsymbol{\pi}_j = [\pi_j^{(0)}, \cdots, \pi_j^{(K-1)}]^T$ represents the current belief about the noise state z_j. When the assumed noise prior is the GM model, then (5.12) reduces to (5.5) and $\boldsymbol{\pi}_j$, now independent of j, is given by the prior impulsive noise model parameters (see Section 5.3.2). However, for the GHMM noise, these beliefs are passed from the MC decoding of the noise state sub-graph; thus, $\pi_j^{(k)} \propto \mu_{z_j \to p_j^i}(z_j = k)$. The resulting input scalar estimation function and its derivative are given by

$$\begin{aligned} g_{\mathsf{in},j}(\hat{r}, \gamma^r) &= \sum_{k=0}^{K-1} \alpha_j^{(k)} \frac{\gamma_k \hat{r}}{\gamma_k + \gamma^r} \\ g'_{\mathsf{in},j}(\hat{r}, \gamma^r) &= \frac{1}{\gamma^i} \Big[-\|g_{\mathsf{in},j}(\hat{r}, \gamma^r)\|^2 \\ &\quad + \sum_{k=0}^{K-1} \frac{\alpha_j^{(k)}}{\gamma^r + \gamma_k} \left(\gamma^r \gamma_k + \frac{\|\gamma_k \hat{r}\|^2}{\gamma^r + \gamma^k} \right) \Big] \end{aligned}$$

where $\gamma_0 = 0$ (absence of impulsive noise) and $\alpha_j^{(k)}$ is the posterior noise state belief given by

$$\alpha_j^{(k)} = P(z_j = k|\hat{r}) = \frac{p(\hat{r}|z_j = k)\, \pi_j^{(k)}}{\sum_{l=0}^{K-1} p(\hat{r}|z_j = l)\, \pi_j^{(l)}} \tag{5.13}$$

where $p(\hat{r}|z_j = k) = \mathcal{N}(\hat{r}; 0, \gamma^r + \gamma_k)$ is the noise state likelihood. Using these input and output scalar estimation functions, GAMP$(\mathbf{i}, \mathsf{l}, \mathbf{F})$ is iterated until convergence generating close approximations of the conditional means $\hat{\mathbf{i}}$ and variances $\boldsymbol{\gamma}^i = [\gamma_0^i, \cdots, \gamma_{N-1}^i]^T$ given the observations $\mathbf{y}$ and the prior information. In the case of GHMM noise, I have additional equalizer iterations where the resulting GAMP messages in the form of noise state likelihoods are passed back to the MC sub-graph of the bursty noise as shown in Figure 5.1. These

state beliefs are given by

$$\mu_{p_j^i \to z_j}(z_j) \propto \mathcal{N}\left(\hat{r}; 0, \gamma^r + \gamma_{z_j}\right).$$

Since the MC sub-graph is non-loopy, only one iteration of forward-backward message passing is needed. This is a standard procedure and, for the interest of space, I refer the reader to [59] and [12] for more details. After that, the resulting noise-state posteriors are passed back to the GAMP algorithm where they are treated as $\boldsymbol{\pi}_j$ in (5.13) in the next equalizer iteration.

When the noise-state likelihoods passed between GAMP$(\mathbf{i}, \mathsf{I}, \mathbf{F})$ and the corresponding MC sub-graph have converged, the equalizer iterations are terminated and messages are passed leftward from the impulse noise sub-graph to the $\mathbf{y}$ factors. Following this, SP rules dictate that the symbol-belief propagating leftward from the y_k node is given by

$$\mu_{y_k \to s_k}(s) = \mathcal{N}\left(y_k; s\hat{\mathsf{H}}_k + \hat{\mathsf{I}}_k, \|s\|^2 \gamma_k^{\mathsf{H}} + \gamma_k^{\mathsf{I}} + \gamma_0\right) \tag{5.14}$$

for all $k \in \mathcal{D}$. (Since null and pilot symbols are known with certainty, there is no need to update their pmfs). Here, $(\hat{\mathsf{H}}_k, \gamma_k^{\mathsf{H}})$ and $(\hat{\mathsf{I}}_k, \gamma_k^{\mathsf{I}})$ play the role of frequency-domain soft channel and impulsive noise estimates, respectively. Furthermore, by SP rules, these messages are then copied leftward to the symbol-mapping nodes so that $\mu_{s_k \to \mathcal{M}_k}(s) = \mu_{y_k \to s_k}(s)$.

Next, I pass the coded-bit beliefs from the symbol-mapping node $\mathcal{M}_k$ to the corresponding bit nodes $c_{k,m}$. Based on SP rules, the messages take the

form

$$\begin{aligned}\mu_{\mathcal{M}_k \to c_{k,m}}(c) &= \sum_{l=1}^{|\mathbb{S}|} \sum_{\mathbf{c}_k \setminus c_m} P\left(s^{(l)} | \mathbf{c}_k\right) \mu_{s_k \to \mathcal{M}_k}(s^{(l)}) \\ &\quad \times \prod_{m' \neq m} \mu_{c_{k,m'} \mathcal{M}_k \to (c_{m'})} \\ &= \frac{\sum_{l: c_m^{(l)} = c} \mu_{s_k \to \mathcal{M}_k}(s^{(l)}) \mu_{\mathcal{M}_k \to s_k}(s^{(l)})}{\mu_{c_{k,m} \to \mathcal{M}_k}(c)}\end{aligned}$$

where the last step is derived in [85].

Finally, the computed coded-bit beliefs are passed to the coding/interleaving factor node. This can be viewed as passing extrinsic soft information in the form of coded-bit priors to the soft-input/soft-output (SISO) decoder. Since SISO decoding has been studied extensively, I refer the interested reader to [59] for a detailed account. As the SISO decoding terminates, it will produce extrinsic soft coded-bit information which will be passed rightward to the symbol-mapping nodes starting a new turbo iteration. The turbo iterations are terminated until the decoder detects no bit errors, the soft bit information has converged, or a maximum number of iterations has been reached.

5.5.4 Simplified Receivers

While the JCNED receiver, as presented, utilizes all the available tones to perform inference over the complete factor graph given in Figure 5.1, the proposed framework is highly flexible providing a trade-off between performance and computational complexity. For example, a receiver, due to computational or architectural constraints, might opt to simplify the receiver by

either (1) restricting the algorithm to a subset $\mathcal{U}$ of the available tones, or (2) simplifying the factor graph structure to avoid passing additional messages.

The selection of the subset $\mathcal{U}$ depends on the desired performance at a given SNR (see Section 5.5.5) and the complexity of the corresponding output estimation functions given in Table 5.2 and Table 5.3. I denote the receiver utilizing only the tones in the subset $\mathcal{U}$ by JCNED $(\mathcal{U})$. The channel and noise estimates on the remaining tones can be obtained by taking the DFT of their corresponding time-domain estimates produced by the GAMP$(\mathbf{h}, \mathsf{H}_{\mathcal{U}}, \sqrt{N}\mathbf{F}_{\mathcal{U}})$ and GAMP$(\mathbf{i}, \mathsf{I}_{\mathcal{U}}, \mathbf{F}_{\mathcal{U}})$ employed by JCNED $(\mathcal{U})$.

On the other hand, simplifying the factor graph can be achieved by using the marginal GM distribution of the GHMM used for modeling clustered channels and bursty noise; thus, the resulting factor graph removes the MC subgraphs corresponding to the noise states and channel support effectively ignoring any dependence within the channel taps and noise samples. While detrimental to performance (see Section 5.6), this avoids computing the messages required by the equalizer iterations. Another approach for simplifying the factor graph is using an independent channel estimate, such as an LMMSE estimator[2] based on pilot tones. Although highly suboptimal in impulsive noise, it eliminates the channel sub-graph from the factor graph and saves the computations required by its corresponding GAMP$(\mathbf{h}, \mathsf{H}_{\mathcal{U}}, \sqrt{N}\mathbf{F}_{\mathcal{U}})$.

[2]In this case, the LMMSE channel estimate and the resulting error will replace the quantities $\hat{\mathsf{H}}_k$ and γ_k^{H} in the scalar impulsive noise estimator of Table 5.2.

5.5.5 Tones Allocation and Selection

The JCNED receiver provides a framework to utilize different tones of an OFDM system to jointly decode data and estimate the channel and the impulsive noise. From this, two questions arise naturally: (i) how should the OFDM system allocate its null, pilot, and data tones?, and (ii) given a fixed number of tones that the receiver is willing to use due to computational constraints, how should it pick the set $\mathcal{U}$ to maximize communication performance? In AWGN channels, it has been shown that a uniformly-spaced placement of pilots is optimal in the MMSE sense [9]. Similarly, null tones are typically allocated at the edge of the spectrum to reduce out of band emissions [1]. However, this no longer holds in impulsive channels as the MMSE channel estimator is no longer linear and the noise is dependent across subcarriers; thus, the information on the null tones can no longer be ignored for data decoding.

The impact of each tone on the ability of the receiver to reconstruct the impulsive noise vector is determined by its type: null tones provide the most information about the impulse noise vector but no information about the channel taps, pilot tones provide information about both the channel taps and the impulsive noise, data tones provide information about both the channel taps and the impulsive noise although not as informative as the null and pilot tones due to the presence of the transmitted data.

If I view the impulsive noise estimation as a sparse reconstruction problem [17], the impact of the tone locations, given by $\mathcal{U}$, on the reconstruction

performance can be evaluated from the properties of the resulting DFT sub-matrix $\mathbf{F}_{\mathcal{U}}$ that is utilized as the measurement matrix Φ in GAMP$(\mathbf{i}, \mathsf{l}_{\mathcal{U}}, \mathbf{F}_{\mathcal{U}})$. A common metric for evaluating the reconstruction performance of a normalized dictionary Φ is the coherence given by

$$\mu_{\Phi} = \max_{k,l,k\neq l} \|\phi_k^* \phi_l\| \tag{5.15}$$

where $\Phi = [\phi_1 \cdots \phi_N]$ (see [89] for recent overview). A simple heuristic for selecting a tone allocation $\mathcal{U}^*$ is given by

$$\mathcal{U}^* = \arg\max_{\mathcal{U}} \ \mu_{\mathbf{F}_{\mathcal{U}}}.$$

While I don't claim any formal optimality of such selection in terms of communication performance, simulation results in Section 5.6 validate using this heuristic in practice.

5.5.6 Computational Complexity

The computational complexity of JCNED depends on two main factors: (i) the number of iterations performed and (ii) the number of utilized tones in the inference given by $\mathcal{U}$. Recall that due to the loopy nature of the system's factor graph, there are two types of iterations: the "turbo iteration" during which messages are passed globally between all the nodes, and the local "equalizer" iterations between the noise and channel dense sub-graphs and their MC sub-graphs in the case of GHMM noise and clustered channels, respectively. Since the MC sub-graphs are non-loopy, they require only one pass of messages ($\mathcal{O}(N)$); thus, the complexity of a single "turbo iteration" is dominated

by the complexity of GAMP($\mathbf{i}, \mathsf{I}_{\mathcal{U}}, \mathbf{F}_{\mathcal{U}}$) and GAMP($\mathbf{h}, \mathsf{H}_{\mathcal{U}}, \sqrt{N}\mathbf{F}_{\mathcal{U}}$) which can be run in parallel. Since both GAMP implementations have $\Phi \propto \mathbf{F}$, the term $\|\Phi_{ij}\|^2$ of Table 5.1 has a constant modulus and steps (R2) and (R6) reduce to a single summation $\forall i$. Furthermore, steps (R1) and (R7) can be implemented efficiently using the FFT algorithm for large $|\mathcal{U}|$ (comparable to the number of tones N) with a complexity $\mathcal{O}(N \log N)$. For smaller $|\mathcal{U}|$, it might be more efficient to implement steps (R1) and (R7) as an matrix-multiply with a complexity $\mathcal{O}(|\mathcal{U}|^2)$. The remaining GAMP operations are scalar operations with complexity $\mathcal{O}(|\mathcal{U}|)$ when implemented sequentially and $\mathcal{O}(1)$ when parallelized. As a result, the total complexity per turbo iteration is $\mathcal{O}(\min\{N \log N, |\mathcal{U}|^2\})$, the same complexity of a typical DFT receiver.

5.6 Numerical Results

In this section, I study the performance of my proposed message-passing JCNED framework by the means of Monte-Carlo simulations. I demonstrate that the proposed JCNED framework provides significant gains (within 1dB from a lower bound) in both coded and uncoded communication systems at a computational complexity slightly higher than the typical DFT receiver and significantly lower than competing prior work. Furthermore, I present numerical results that provide insights into more fundamental questions concerning the value of noise modeling, the value of joint data decoding and channel/noise estimation, and the impact of the number of tones and their allocation on communication performance.

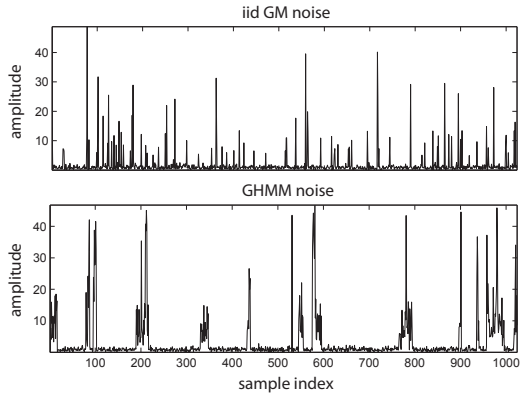


Figure 5.3: The amplitude of two noise traces with the same sample marginals but different temporal dynamics: in the iid model samples are generated independently, while the GHMM exhibits bursty behavior.

5.6.1 Setup

In my simulations, I consider either $N = 256$ (narrowband PLC systems) or 1024 (modern wireless systems) subcarriers modulated by either 4-QAM or 16-QAM constellations. I assume a dispersive Rayleigh fading channel with 10 taps for wireless systems and 5 taps for PLC[3]. Unless stated otherwise, pilot tones are assigned uniformly while the null tones are chosen randomly. The realizations of the impulsive noise were generated from two models: an iid GM model having two impulsive components with powers 20dB and 30dB above the background noise occurring 7% and 3% or the time, respectively; and an GHMM having the same marginal distribution and impulsive power

[3]Due to space limitations, I choose to focus on a Rayleigh channel (although its estimation is still highly non-linear), rather than the clustered-tap channel whose inference is detailed in [85].

with a state transition matrix given by

$$\mathbf{T} = \begin{bmatrix} 0.989 & 0.006 & 0.005 \\ 0.064 & 0.857 & 0.079 \\ 0.183 & 0.150 & 0.667 \end{bmatrix}.$$

Even though both models have the same marginal, their traces, shown in Figure 5.3, have very different realizations due to their temporal dependencies. Unless stated otherwise, the results have been generated using 5 turbo iterations, and 15 GAMP iterations. In all results, the signal to noise ratio (SNR) refers to the ratio of the signal power to the second order moment of the noise.

5.6.2 Comparison With Other Schemes

Figure 5.4 shows the uncoded symbol error rate (SER) comparison of my proposed JCNED framework to the prior work discussed in Section 2.4 for a 4-QAM modulated OFDM system with 256 subcarriers of which 80 are null tones and 15 are pilots under a 5-tap Rayleigh channel corrupted by iid GM noise. In addition to the typical OFDM receiver, labeled "DFT", that takes the DFT of the received signal and performs independent detection on each subcarrier, I choose to compare against the MMSE receiver [45] and the SBL receiver [57], labeled "MMSE" and "SBL", respectively[4]. The MMSE receiver has the best performance among the "time-domain preprocessing techniques" since it is optimal in the MMSE sense when the temporal dependence in the OFDM signal is ignored, an assumption shared among these techniques. Similarly, the SBL receiver was shown to have the best performance among the

[4]I have augmented the MMSE and SBL receivers to perform channel estimation using an LMSSE estimator, something missing in the original formulation in [45, 57]

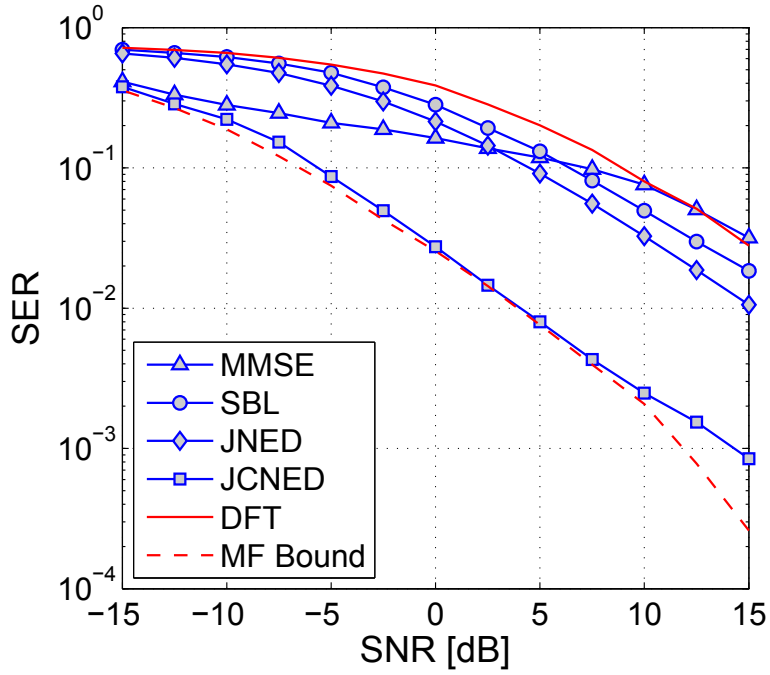


Figure 5.4: Uncoded SER of different schemes for 4-QAM and $N = 256$ subcarriers with $N_{\mathsf{n}} = 80$ null subcarriers and $N_{\mathsf{p}} = 15$ pilot subcarriers.

"Sparse Impulse Noise Reconstruction" methods [57]. The SER curves in Figure 5.4 show that the proposed JCNED receiver outperforms the DFT receiver by 15dB, the SBL receiver by 11dB, and the MMSE receiver by 7dB in the low SNR regime and by 15dB in the high SNR regime. This huge performance gain is due to JCNED utilizing all the tones for impulsive noise and channel estimation as opposed to SBL and MMSE utilizing only the null and pilot tones and using a LMMSE channel estimator. To further illustrate the effect of using all available tones and the impact of joint impulsive noise and channel estimation, I plot the SER curve corresponding to a simplified version of JCNED $(\mathcal{N} \cup \mathcal{P})$that uses only the known tones and the LMMSE estimate of

the channel instead of jointly estimating it. I denote this receiver as a joint noise estimation and data decoding (JNED). From Figure 5.4, JNED is only 2.5dB better than SBL and worse than MMSE in the low SNR regime.

The MMSE has the lowest computational complexity, followed by JCNED, and then SBL. This is reflected in the simulations' running time with MMSE being the fastest. Since they utilize GAMP, the JCNED and the JNED receivers had a comparable running time to the MMSE receiver as expected from the discussion in Section 5.5.6. On the other hand, the SBL receiver, requiring an huge matrix inversion, was by far the slowest with running time 100x slower than JCNED.

A lower bound on the communication performance of OFDM systems in impulsive noise can be obtained by considering the transmission of a symbol on a single data tone while nulling all the remaining tones and having perfect channel information. I refer to this bound as the *matched filter* bound. Referring to (5.1), the received signal corresponding to sending a symbol s on tone k is given by

$$\mathbf{r} = \mathbf{H}\mathbf{F}^*(s\mathbf{e}_k) + \mathbf{n} = s\bar{\mathbf{f}}_k + \mathbf{n} \tag{5.16}$$

where $\mathbf{e}_k$ is the standard basis and $\bar{\mathbf{f}}_k$ is the k-th column of $\mathbf{H}\mathbf{F}^*$. Eq. (5.16) represents an N-branch diversity under impulsive noise which was analyzed in [88] for the case of a unit flat fading channel. Figure 5.4 shows the SER attained by the MF bound labeled as "MF Bound". It can be seen that my proposed JCNED receiver follows the MF bound closely to within 1dB with a slight deviation at very high SNR which I suspect is to the high nonlinearity

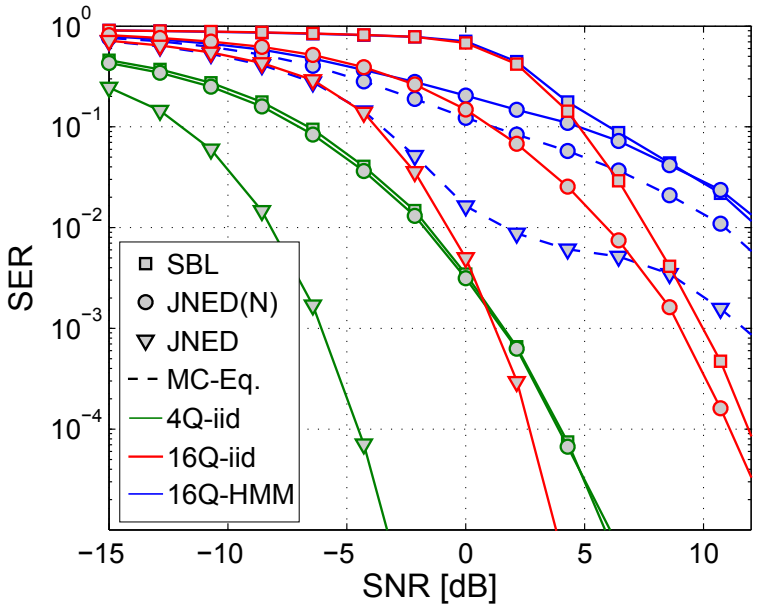


Figure 5.5: Uncoded SER in a flat unit channel for an OFDM system of $N = 256$ subcarriers with $N_{\mathsf{n}} = 60$ null tones.

of the output channel for GAMP$(\mathbf{i}, \mathsf{I}_{\mathcal{U}}, \mathbf{F}_{\mathcal{U}})$ (at this SNR there is virtually no background noise).

5.6.3 The Value of Impulsive Noise Modeling

In this section, I quantify the benefit of training a noise model vs. treating it as a sparse reconstruction problem as done by the SBL receiver. Figure 5.5 shows the SER performance of my receiver framework (labeled JNED since there is no channel to estimate) compared to SBL for a unit flat fading channel in both the iid GM and GHMM noise. I chose a unit flat fading channel to isolate the effect of noise modeling on communication performance. For each setting, I compare SBL against JNED $(\mathcal{N})$, that uses only the null

tones, and against JNED that uses all available tones. The performance of JNED ($\mathcal{N}$) serves as a fair evaluation for the value of the noise model since SBL doesn't make use of the remaining data tones. There are two sets of SER curves for the iid GM noise: one for 4-QAM modulation marked in green, and another for 16-QAM modulation marked in red. Figure 5.5 shows that the performance of JNED ($\mathcal{N}$) is very close to that of SBL for 4-QAM modulation. This small performance difference is due to the fact that detection of 4-QAM symbols depends only on the phase of the estimated noise and is indifferent to the accuracy of the noise amplitude which is better in JNED. However, as seen in Sections 5.5.6 and 5.6.2, there is a huge computational and architectural (order of magnitude faster and can be parallelized) advantage to using the JNED algorithm over SBL.

Figure 5.6 plots the normalized mean squared error (NMSE) of the noise estimate under each method. The noise NMSE is given by $\|\mathbf{n} - \hat{\mathbf{i}}\|^2/\gamma^n$; thus, for a unit power signal MSE = NMSE + SNR . JNED always results in a lower NMSE than SBL with significant MSE difference across a wide SNR range. As a result, I expect significant improvement in performance when employing higher order modulations as shown in Figure 5.5 for 16-QAM marked in red (up to 6dB). Furthermore, using all the tones increases the performance by an additional 8-10dB for both modulations. This is explained by the significantly lower NMSE in Figure 5.6 when using all the tones. Interestingly, when the estimation uses only the null tones, the NMSE is independent of the modulation order; however, when utilizing all the tones, the NMSE for 16-

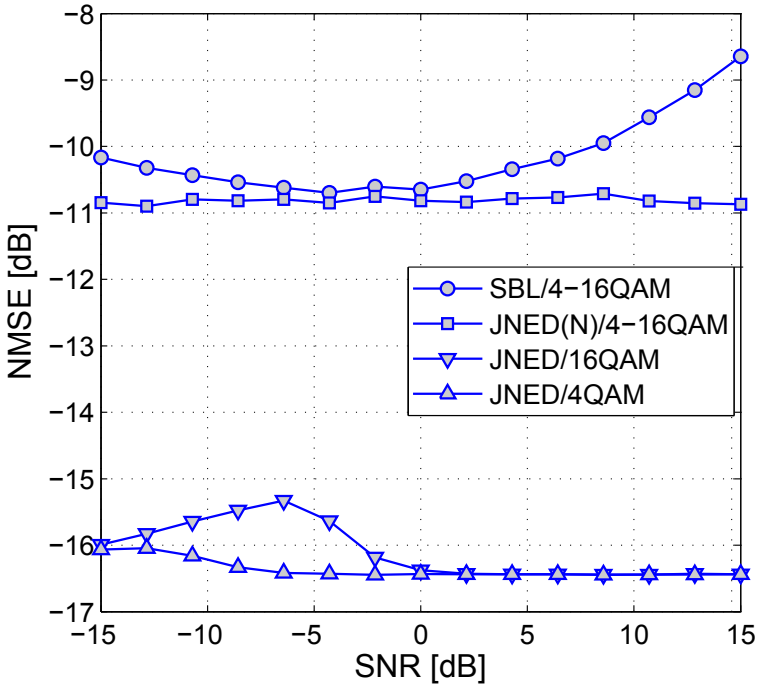


Figure 5.6: Normalized MSE of the noise estimates produced by JNED and SBL with SER performance given in Figure 5.5.

QAM modulation is higher than that of 4-QAM for a moderate range of SNR . This can be explained by the fact that it is harder to distinguish a 16-QAM modulated signal than the constant amplitude 4-QAM signal in the presence of impulses. For GHMM noise, denoted in blue, I consider two categories of receivers: (i) simplified receivers based on the GHMM noise model marginal that ignore the temporal dependence of the noise to avoid MC equalization represented by solid lines, and (ii) receivers with MC equalization. Similar to the marginal JNED, SBL doesn't make use of the temporal dependency in the noise samples; however, the marginal JNED receiver is 3-6dB better than

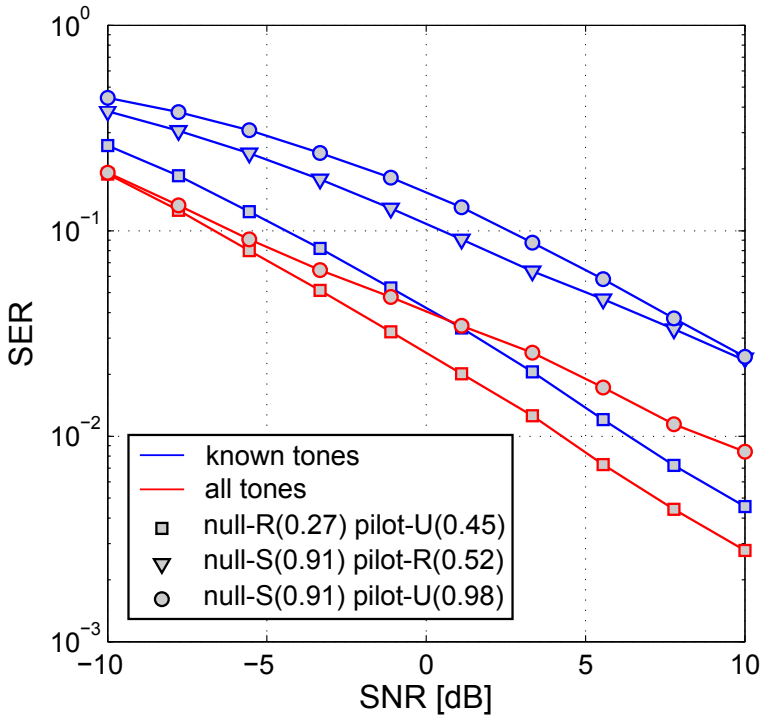


Figure 5.7: Uncoded SER in iid GM noise under a 5-tap Rayleigh channel for an OFDM system with 256 tones with 60 null tones and 25 pilots allocated in different configurations.

the SBL receiver. The JNED ($\mathcal{N}$) receiver employing MC equalization adds an additional 2.5dB in performance gain over the marginal JNED while using all tones adds another 8dB.

5.6.4 Known Tone Allocation

Section 5.5.5 discussed the impact of known, either pilot or null, tones on impulsive noise reconstruction and communication performance. Figure 5.7 shows the uncoded SER performance for JCNED in iid GM noise under a 5-

tap Rayleigh channel. In this figure, blue indicates running JCNED on known tones while red indicates running it with all tones. Furthermore, "*type*-$X(\mu)$" indicates that a tone of *type* either a null or a pilot has an allocation pattern X (R-random, U-uniform, S-sideband) resulting in a dictionary coherence given by μ. Figure 5.7 shows that the typical allocation of nulls at the sides of the band and spreading the pilots uniformly produces the worst performance. Randomizing the pilot tones alone improves the performance by 2.5dB while randomizing the null tones improves the performance by 7dB. This also holds when utilizing all the tones with a performance gain of around 5dB when randomizing the null tone allocation. In addition, Figure 5.7 shows that the lower the coherence of a certain allocation the better the communication performance. This result opens new interesting research directions for known tone allocation in impulsive noise channels where conventional allocation techniques for the AWGN channel fail to account for impulsive noise estimation accuracy.

5.6.5 Coded Systems

Figure 5.8 shows the bit error rate (BER) performance of a coded OFDM system with 1024 subcarriers with 150 pilots in a 10-tap Rayleigh channel corrupted by iid GM noise. Due to the aggressive nature of impulsive channels, I employ LDPC codes with code-word length $\sim 60\,000$ and rate 1/2. The label "alg-#" refers to the algorithm used by the receiver followed by the number of turbo iterations performed. I consider three types of receivers: the typical receiver that takes the DFT of the received signal and performs sym-

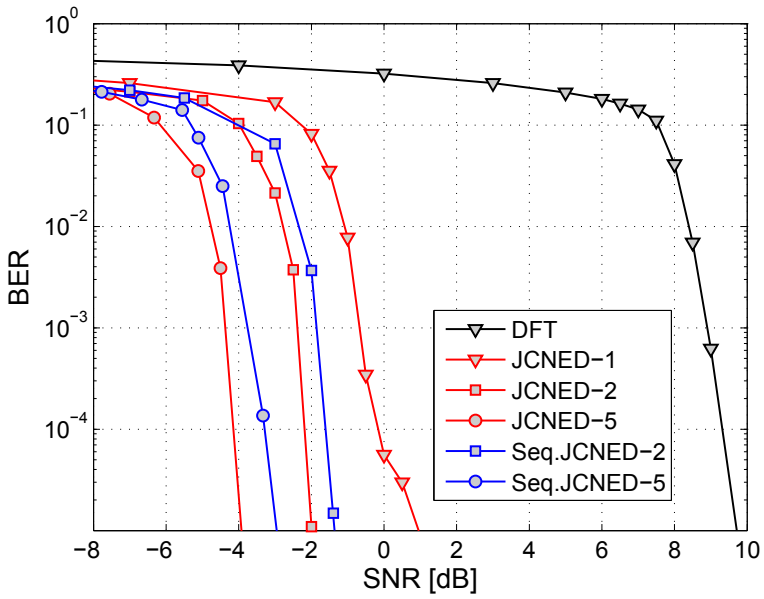


Figure 5.8: LDPC coded BER in iid GM noise under a 10-tap Rayleigh channel for an OFDM system with 1024 tones with 150 pilots.

bol detection followed by LDPC decoding; the proposed JCNED receiver that performs joint channel/noise estimation and LDPC decoding; and a simplified version of JCNED, labeled "Seq. JCNED", that first performs channel/noise estimation and symbol detection as in uncoded systems delaying the LDPC decoding till the last turbo iteration. In all receivers, I set the maximum LDPC iterations to 50. With only 1 turbo iteration, JCNED[5] provides an additional 9dB over the coded DFT receiver. An additional turbo iteration yields an extra 2dB while 5 turbo iterations in total yield a 13dB improvement over the DFT receiver. Furthermore, by decoding the LDPC code in each turbo

[5]Since there is only one turbo iteration, JCNED-1 and seq. JCNED-1 are identical.

iteration, JCNED provides an additional 1dB over seq. JCNED.

5.6.6 Data Rates

Section 5.6.5 provided BER simulations achieved by the proposed receiver. In practice, information bits are coded and divided into data packets for transmission. A collection of erroneous bits in a received packet might render all the information in the packet useless. As a result, BER by itself is not enough to characterize the data rates achieved by a given receiver.

The transmission rate R of an OFDM system can be related to its bandwidth B and modulation order M (assuming same order across all subcarriers) as $R = \eta N_{\mathsf{d}} \log_2(M) \times B/(N + L_c)$, where η is the code rate, N is the number of tones, N_{d} is the number of data tones, and L_c is the length of the cyclic prefix. For simplicity, let us assume that the information bits can be mapped to one physical layer packet. Then, ignoring overhead at the MAC layer, the rate of our system can be written as $R = M_{\mathsf{i}} B/Q(N + L_c)$ where M_{i} is the number of information bits at the application level and Q is the number of OFDM symbols in a physical layer packet. However, due to packet losses, the communication system will operate below this nominal rate.

Goodput characterizes data throughput from the perspective of the number of information bits *successfully* transmitted to the receiver per unit time. The goodput G is defined as $G = R(1 - P_e)$ where P_e is the probability of physical layer packet error (also data packet based on our assumption) [4, 5, 100]. The probability of success $P_s = 1 - P_e$ represents the fraction of

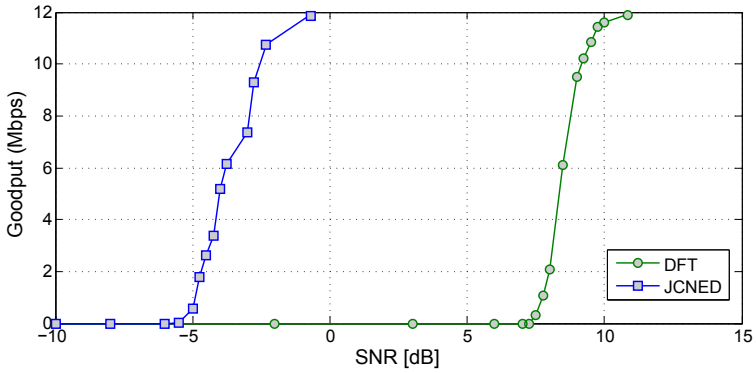


Figure 5.9: Goodput vs SNR for packets of size 32400bits: The proposed JCNED receiver achieves the same rate at a 12dB lower SNR than the traditional DFT receiver.

successful packet transmissions. A different interpretation of G could be seen by treating the data packet transmission and retransmission as a Bernoulli trial with success probability P_s. Then, the expected number of packet retransmissions is given by $1/P_s$. Since this number of retransmissions is needed for the same packet the effective rate is then $R \times P_s$ which is the definition of goodput G.

For simulations, I used the same system parameters as in Section 5.6.5 and chose $B = 10$MHz, and $Q(N + L_c) = 142.7\mu$sec (inspired by LTE). For these settings, each data packet (32400bits) is mapped to a physical layer packet made of 19 OFDM symbols. As a result, $R = 32400/(19 \times 142.7) \approx$ 12Mbps. Figure 5.9 shows the goodput vs SNR for these system settings. The proposed receiver achieves a gain of about 12db over the DFT receiver allowing the receiver to operate in much higher interference environments. A similar result is shown in Figure 5.10 that shows goodput vs interference power at very

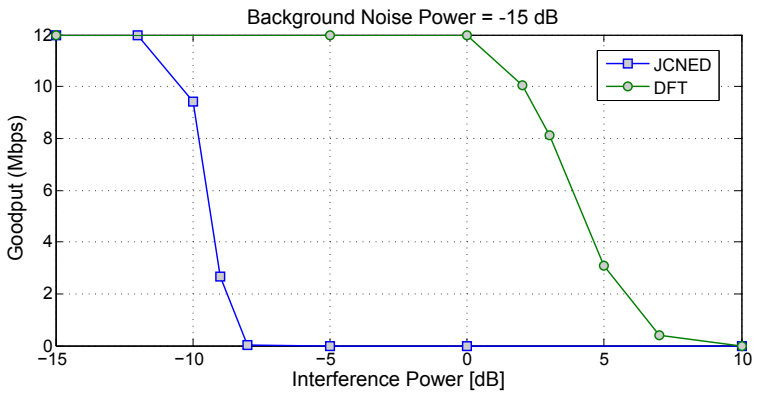


Figure 5.10: Goodput vs interference power for packets of size 32400bits and background power of −15dB: The proposed JCNED receiver achieves the same rate at a 13dB higher interference power than the traditional DFT receiver.

low background noise power. Again our receiver is able to tolerate interference power that is 13dB higher than those tolerated by the DFT receiver.

5.7 Conclusion

In a this chapter, I presented a factor-graph approach to joint channel/noise estimation and data decoding in impulsive noise channels. My approach merges recent work on modeling impulsive noise in communication systems [41] with recent advances in approximate message passing algorithms [28, 79, 80] and SISO decoding [59]. The presented receiver has a complexity comparable to the typical DFT receiver while providing tens of dBs in performance gain (1dB from a lower bound). Furthermore, it can be parallelized which provides a natural mapping to FPGA implementations. In addition, my experiments addressed more fundamental questions such as the value of

impulsive noise modeling, and null and pilot allocation in impulsive noise. The former provides new interesting ideas for future research.

In the following chapter, I address the question of noise parameter estimation without an explicit training interval. This will make the proposed receivers of this chapter more robust to fast varying environments.

5.8 Appendix

5.8.1 Derivation of GAMP$(\mathbf{h}, \mathsf{H}_{\mathcal{U}}, \sqrt{N}\mathbf{F}_{\mathcal{U}})$ Functions

This appendix describes the derivation of the channel output scalar estimation functions $g_{\mathsf{out},k}(y_k, \hat{\mathsf{H}}, \gamma^{\mathsf{H}})$ and $-g'_{\mathsf{out},k}(y_k, \hat{\mathsf{H}}, \gamma^{\mathsf{H}})$ for data tones. The derivation for pilot tones is relatively straightforward and follows from linear estimation theory [53]. The data tone output channel is given by

$$y_k = s_k \mathsf{H}_k + \mathsf{I}_k + \mathsf{G}_k \qquad k \in \mathcal{D}.$$

Given s_k, this channel is Gaussian and

$$\mathbb{E}\left\{\mathsf{H}_k | s_k, y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\} = \hat{\mathsf{H}} + \frac{s_k^* \gamma^{\mathsf{H}} \left(y_k - \hat{\mathsf{I}}_k - s_k \hat{\mathsf{H}}\right)}{\gamma_0 + \gamma_k^{\mathsf{I}} + \|s_k\|^2 \gamma^{\mathsf{H}}}. \tag{5.17}$$

Given the belief β_k about symbol s_k, we use the *law of total expectation* to obtain

$$\begin{aligned}\mathbb{E}\left\{\mathsf{H}_k | y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\} &= \mathbb{E}_{s_k|y_k}\left\{\mathbb{E}\left\{\mathsf{H}_k | s_k, y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}\right\} \\ &= \hat{\mathsf{H}} + \sum_{l=1}^{|\mathbb{S}|} \lambda_k^{(l)} \frac{\gamma^{\mathsf{I}} \left(y_k - \hat{\mathsf{I}}_k - \hat{\mathsf{H}} s^{(l)}\right)}{\gamma_0 + \gamma_k^{\mathsf{I}} + \|s^{(l)}\|^2 \gamma^{\mathsf{H}}}\end{aligned} \tag{5.18}$$

where the posterior symbol belief is given by

$$\lambda_k^{(l)} = P\left(s_k = s^{(l)}|y_k\right) \propto p\left(y_k|s_k\right)\beta_k^{(l)}$$

and $p\left(y_k|s_k\right) = \mathcal{N}\left(y_k; \hat{\mathsf{I}}_k + \hat{\mathsf{H}}s_k, \gamma^{\mathsf{I}} + \|s_k\|^2\gamma^{\mathsf{H}} + \gamma_0\right)$. Similarly, using the *law of total variance*, for $\forall k \in \mathcal{D}$ we have

$$\begin{aligned}
&\mathbb{V}\left\{\mathsf{H}_k|y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\} \\
&\quad = \mathbb{E}_{s_k|y_k}\left\{\mathbb{V}\left\{\mathsf{H}_k|s_k, y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}\right\} \\
&\qquad + \mathbb{V}_{s_k|y_k}\left\{\mathbb{E}\left\{\mathsf{H}_k|s_k, y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}\right\} \\
&\quad = \sum_{l=1}^{|\mathbb{S}|} \lambda_k^{(l)}\left[\frac{\gamma^{\mathsf{H}}\left(\gamma_0 + \gamma_k^{\mathsf{I}}\right)}{\gamma_0 + \gamma_k^{\mathsf{I}} + \|s^{(l)}\|^2\gamma^{\mathsf{H}}}\right. \\
&\qquad \left. + \left\|\mathbb{E}\left\{\mathsf{H}_k|s_k, y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}\right\|^2\right] - \left\|\mathbb{E}\left\{\mathsf{H}_k|y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}\right\|^2
\end{aligned}$$

where $\mathbb{E}\left\{\mathsf{H}_k|s_k, y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}$ and $\mathbb{E}\left\{\mathsf{H}_k|y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}$ are given in (5.17) and (5.18), respectively. The resulting expressions for $\mathbb{E}\left\{\mathsf{H}_k|y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}$ and $\mathbb{V}\left\{\mathsf{H}_k|y_k; \hat{\mathsf{H}}, \gamma^{\mathsf{H}}\right\}$ can be plugged into steps (D2-3) of Table 5.1 to arrive at the desired result given in Table 5.2.

5.8.2 Derivation of GAMP$(\mathbf{i}, \mathsf{I}_{\mathcal{U}}, \mathbf{F}_{\mathcal{U}})$ Functions

This appendix describes the derivation of the impulsive noise output scalar estimation functions $g_{\mathsf{out},k}(y_k, \hat{\mathsf{I}}, \gamma^{\mathsf{I}})$ and $g'_{\mathsf{out},k}(y_k, \hat{\mathsf{I}}, \gamma^{\mathsf{I}})$ for data tones. The derivation for null and pilot tones is relatively straightforward and follows from linear estimation theory [53]. The data tone output channel is given by

$$y_k = \mathsf{I}_k + s_k\mathsf{H}_k + \mathsf{G}_k \qquad k \in \mathcal{D}.$$

Given s_k, this channel is Gaussian and

$$\mathbb{E}\left\{\mathsf{I}_k|s_k, y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\} = \hat{\mathsf{I}} + \frac{\gamma^{\mathsf{I}}\left(y_k - \hat{\mathsf{I}} - \hat{\mathsf{H}}_k s_k\right)}{\gamma_0 + \gamma^{\mathsf{I}} + \|s_k\|^2\gamma_k^{\mathsf{H}}}. \tag{5.19}$$

Given the belief β_k about symbol s_k, we use the *law of total expectation* to obtain

$$\begin{aligned}\mathbb{E}\left\{\mathsf{I}_k|y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\} &= \mathbb{E}_{s_k|y_k}\left\{\mathbb{E}\left\{\mathsf{I}_k|s_k, y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}\right\} \\ &= \hat{\mathsf{I}} + \sum_{l=1}^{|\mathbb{S}|} \lambda_k^{(l)} \frac{\gamma^{\mathsf{I}}\left(y_k - \hat{\mathsf{I}} - \hat{\mathsf{H}}_k s^{(l)}\right)}{\gamma_0 + \gamma^{\mathsf{I}} + \|s^{(l)}\|^2\gamma_k^{\mathsf{H}}}\end{aligned} \tag{5.20}$$

where the posterior symbol belief is given by

$$\lambda_k^{(l)} = P\left(s_k = s^{(l)}|y_k\right) \propto p\left(y_k|s_k\right)\beta_k^{(l)}$$

and $p\left(y_k|s_k\right) = \mathcal{N}\left(y_k; \hat{\mathsf{I}} + \hat{\mathsf{H}}_k s_k, \gamma^{\mathsf{I}} + \|s_k\|^2\gamma_k^{\mathsf{H}} + \gamma_0\right)$. Similarly, using the *law of total variance*, for $\forall k \in \mathcal{D}$ we have

$$\begin{aligned}&\mathbb{V}\left\{\mathsf{I}_k|y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\} \\ &= \mathbb{E}_{s_k|y_k}\left\{\mathbb{V}\left\{\mathsf{I}_k|s_k, y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}\right\} \\ &\quad + \mathbb{V}_{s_k|y_k}\left\{\mathbb{E}\left\{\mathsf{I}_k|s_k, y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}\right\} \\ &= \sum_{l=1}^{|\mathbb{S}|} \lambda_k^{(l)}\left[\frac{\gamma^{\mathsf{I}}\left(\gamma_0 + \|s^{(l)}\|^2\gamma_k^{\mathsf{H}}\right)}{\gamma_0 + \gamma^{\mathsf{I}} + \|s^{(l)}\|^2\gamma_k^{\mathsf{H}}}\right. \\ &\quad \left. + \left\|\mathbb{E}\left\{\mathsf{I}_k|s_k, y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}\right\|^2\right] - \left\|\mathbb{E}\left\{\mathsf{I}_k|y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}\right\|^2\end{aligned}$$

where $\mathbb{E}\left\{\mathsf{I}_k|s_k, y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}$ and $\mathbb{E}\left\{\mathsf{I}_k|y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}$ are given in (5.19) and (5.20), respectively. The resulting expressions for $\mathbb{E}\left\{\mathsf{I}_k|y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}$ and $\mathbb{V}\left\{\mathsf{I}_k|y_k; \hat{\mathsf{I}}, \gamma^{\mathsf{I}}\right\}$ can be plugged into steps (D2-3) of Table 5.1 to arrive at the desired result given in Table 5.2.

Chapter 6

Robust Message-Passing OFDM Receivers for Impulsive Noise Channels

The message-passing receivers proposed in the previous chapter provide significant gains in communication performance (more than 10dB) over the conventional DFT receiver by utilizing the interference models proposed in Chapter 3. However, these receivers assume knowledge of the interference model parameters. While a reasonable assumption when the interference statistics are slowly varying and could be estimated during a quiet period, this will no longer be the case in rapidly changing environments where the interference statistics could change on the order of an OFDM symbol duration. Under such scenarios, the performance of the message-passing receivers might suffer due to *model mismatch* in the assumed interference model. In this chapter, I extend the graphical model framework, proposed in Chapter 5, to perform parameter estimation of the interference model. In addition, I propose replacing the interference model priors by the *automatic relevance determination* (ARD) prior that can take advantage of temporal sparsity commonly present in the interference using an empirical Bayesian framework called *sparse Bayesian learning* (SBL).

6.1 Introduction

As discussed in previous chapters, interference in the form of impulsive noise has a detrimental effect on communication performance. OFDM modulation provides resilience to this interference by coding the data symbols over multiple time samples. While the conventional DFT receiver doesn't realize the expected gain in communication performance [46, 47], the message-passing receivers proposed in Chapter 5 provide huge performance gains (tens of dBs). These receivers leverage accurate parametric models of the interference statistics to perform Bayesian inference of the transmitted data. In particular, they assume perfect knowledge of the interference model parameters. While a valid assumption in slowly varying environments, this might no longer hold if the interference statistics are rapidly changing. This could lead to model mismatch that can deteriorate the communication performance significantly. As a result, it is of interest to design robust OFDM receivers that can adapt to the changing interference environment.

The prior work dealing with robust receivers for impulsive noise channels can be roughly classified into two categories: (i) robust-metric based, and (ii) sparse signal recovery. Robust metrics are designed to minimize the effect of outliers and are typically employed as part of pre-processing mitigation techniques as described in Section 2.4. Two examples are the Huber metric [20] and the modified-soft-limiting metric [30] that has been used for turbo decoding in single carrier systems. A more recent work proposes the use of the correntropy induced metric combined with zero order statistics for prefilter-

ing applications [42]. Receivers based on sparse signal recovery typically treat the impulsive noise vector as an unknown and attempt to reconstruct it using known data tones. These techniques have been discussed in Section 2.4.

6.2 Contribution

In this chapter, I propose the design of robust message-passing OFDM receivers that can adapt to rapidly varying environments. In particular, I propose two classes of robust receivers: (i) blind interference model parameter estimation based, and (ii) automatic relevance determination (ARD) prior based. On one hand, the blind receiver, discussed in Section 6.4, jointly estimates the interference model parameters without any explicit training period while performing joint channel/interference estimation and data decoding. On the other hand, by using the ARD prior in Section 6.5 instead of the GM interference model, I propose a message-passing version of the sparse Bayesian learning (SBL) receiver that achieves excellent performance over a variety of different interference models [57]. However, the message-passing SBL receiver has the following key advantages over the direct SBL implementation: 1) its computational complexity is of the same order as the conventional OFDM receiver; 2) doesn't require any matrix inversions and can be naturally mapped to an FPGA implementation; and 3) it is part of a Bayesian inference framework that jointly estimates the channel and the interference while approximating MAP symbol detection (direct SBL uses the suboptimal LMMSE estimate of the channel and targets minimizing the MSE error of the interference estimate

[57]). For the ease of presentation, this chapter derives the receivers for an uncoded OFDM system; however, it is straightforward to generalize the receivers to the LDPC coded framework presented in the previous chapter.

6.3 System Model

Since the proposed robust receivers are extensions of the message passing receivers described in Chapter 5 they will follow the same system model. However, for ease of presentation and without loss of generality, I focus on the uncoded OFDM system model. The extension of these algorithms to the coded OFDM system of Chapter 5 is straightforward.

Same as in Section 5.3, I consider an N-tone OFDM system with the following tone partition: N_{p} pilot tones indexed by the set $\mathcal{P}$, N_{n} null tones indexed by $\mathcal{N}$, and N_{d} data tones indexed by the set $\mathcal{D}$ where each data subcarrier is modulated by a symbol from an 2^M-ary constellation $\mathbb{S}$. The source-generated data bits are mapped to the data symbols using gray mapping. In the sequel, I use $s^{(i)} \in \mathbb{S}$ for $i \in \left\{1, \cdots, 2^M\right\}$ to denote the ith element of $\mathbb{S}$. Likewise, I use $s_k[q]$ to denote the symbol transmitted on the kth subcarrier of the qth OFDM symbol. Based on the tone partition, I note that: $s_k[q] = \mathsf{p}$ for all $k \in \mathcal{P}$, where p is a known pilot symbol; $s_k[q] = 0$ for all $k \in \mathcal{N}$; and $s_k[q] = s^{(l)}$ for some l. On the frame level, I use $\mathbf{c}[q]$ to denote the coded/interleaved bits allocated to the data tones of the qth OFDM symbol.

As discussed in Section 2.1 and Section 5.3, OFDM modulation applies an N-point inverse discrete Fourier transform (IDFT) $\mathbf{F}^*$ to a vector of N

symbols. After OFDM modulation, the resulting time-domain signal corresponding to the qth OFDM symbol is given by $\mathbf{u}[q] = \mathbf{F}^*\mathbf{s}[q]$ where $\mathbf{u}[q] = [u_0[q] \cdots u_{N-1}[q]]^T$ is the time domain signal and $\mathbf{s}[q] = [s_0[q] \cdots s_{N-1}[q]]^T$ is the transmitted symbol sequence.

After appending the cyclic prefix, the qth OFDM symbol's waveform propagates through a noisy LTI channel with a channel impulse response $\mathbf{h}[q] = [h_0[q] \cdots h_{L-1}[q]]^T$ where L is the number of channel taps. After discarding the cyclic prefix and assuming perfect synchronization at the receiver, the received signal $\mathbf{r}[q]$ can be expressed as

$$\mathbf{r}[q] = \mathbf{H}[q]\mathbf{u}[q] + \mathbf{n}[q] = \mathbf{H}[q]\mathbf{F}^*\mathbf{s}[q] + \mathbf{n}[q] \tag{6.1}$$

where $\mathbf{n}[q]$ is additive noise, and $\mathbf{H}[q]$ is the circulant matrix formed by $\mathbf{h}[q]$ [94]. The receiver applies DFT to $\mathbf{r}[q]$, and the resulting frequency-domain signal is given by

$$\mathbf{y}[q] = \mathbf{F}\mathbf{r}[q] = \mathbf{F}\mathbf{H}[q]\mathbf{F}^*\mathbf{s}[q] + \mathbf{F}\mathbf{n}[q] = \mathsf{H}[q] \circ \mathbf{s}[q] + \mathsf{N}[q] \tag{6.2}$$

where $\mathsf{H}[q] = (\sqrt{N}\mathbf{F}_{:,1:L})\mathbf{h}[q]$ denotes the frequency-domain channel, $\mathsf{N}[q] = \mathbf{F}\mathbf{n}[q]$ the frequency-domain noise, and $\circ$ denotes the Hadamard product.

In this chapter, I assume the GM model of interference as described in Chapter 2. As a result, the additive noise term $\mathbf{n}[q]$ will be a GM distributed random vector. The pdf of a GM distributed random variable n with zero mean components is given by

$$p(n) = \sum_{k=0}^{K-1} \pi^{(k)} \cdot \mathcal{N}(n; 0, \gamma_k) \tag{6.3}$$

where $\mathcal{N}(n; 0, \gamma_k)$ denotes the complex Gaussian pdf with zero mean and variance γ_k; and $\pi^{(k)}$ is the probability of the k-th Gaussian component. Generally, the Gaussian background noise (also called thermal noise) $g \sim \mathcal{N}(0, \gamma_0)$ is the component with the smallest variance that is typically assumed to be known. The total noise can then be decomposed as the sum of background and impulsive noise as $n = g + i$, where i is the impulsive noise component whose pdf is given by

$$p(i) = \pi^{(0)}\delta(i) + \sum_{k=1}^{K-1} \pi^{(k)} \cdot \mathcal{N}(i; 0, \gamma_k). \tag{6.4}$$

Here, $\pi^{(0)}$ represents the probability that only background noise is present; thus, with probability $1 - \pi^{(0)}$ a non-zero impulsive noise is present on top of the background noise. $\left\{\pi^{(k)}\right\}_{k=0}^{K-1}$ and $\{\gamma_k\}_{k=1}^{K-1}$ are parameters to be estimated. When the interference statistics are slowly varying, these parameters can be estimated during a quiet period with no transmission and used in the message passing receivers proposed in Chapter 5. However, in rapidly varying environments, these parameters have to be estimated for every OFDM symbol: a constraint that motivates the receivers proposed in this chapter.

6.4 Joint Channel/Noise Estimation and Decoding with Blind Interference Model Parameter Estimation (Blind JCNED)

In this section, I extend the framework used in Chapter 5 to design OFDM receivers that estimate the interference model parameters without the need for a explicit training period in addition to performing the joint chan-

nel/interference estimation and data decoding.

6.4.1 Revisiting Decoding in Impulsive Noise Channels

As described in Section 5.4, the objective is to infer the information bits $\mathbf{b}$ transmitted on the data tones given the received signal $\mathbf{y}$, the prior knowledge about the impulsive noise and the channel, and the locations of the pilot and null tones given by $\mathcal{P}$ and $\mathcal{N}$, respectively. In Chapter 5, the prior knowledge about the interference was given by the parametric model of the impulsive noise $p\left(i;\theta_{\mathbf{i}}\right)$ and the true value of the parameters given by $\theta_{\mathbf{i}} = \{\gamma_k\}_{k=0}^{K-1}$. In this section, the prior knowledge about the interference is restricted to the parametric model of the impulsive noise $p\left(i;\theta_{\mathbf{i}}\right)$ without knowing the exact value of the parameters $\theta_{\mathbf{i}}$.

A fully Bayesian approach first marginalizes over the channel taps $\mathbf{h}$, the noise vector $\mathbf{n}$ (both considered latent variables), and the impulsive noise parameters $\theta_{\mathbf{i}}$ for some prior $p\left(\theta_{\mathbf{i}}\right)$. Then, it computes the posterior probability of the symbol s_k given the received signal $\mathbf{y}$. Writing the parameter marginalization step explicitly, the posterior probability

$$P\left(s_k|\mathbf{y}\right) = \int_{\theta_{\mathbf{i}}} P\left(s_k|\mathbf{y},\theta_{\mathbf{i}}\right) p\left(\theta_{\mathbf{i}}\right)\, \mathrm{d}\theta_{\mathbf{i}}, \quad \forall k \in \mathcal{D} \tag{6.5}$$

where I omitted the channel parameters for clarity of presentation (these are explicitly considered in Section 5.4). Typically, this marginalization is not readily available in closed form and requires using sampling techniques for its evaluation. However, this is typically computationally intensive and not

appropriate for receiver design. Instead, I opt to use a point estimate of the interference parameters $\hat{\theta}_{\mathbf{i}}$. In contrast, when perfect knowledge of the parameters is readily available as in Chapter 5, we have $p\left(\theta_{\mathbf{i}}\right) = \delta(\theta_{\mathbf{i}} - \theta_{\mathbf{i}}^{t})$ where $\theta_{\mathbf{i}}^{t}$ are the true parameters. Based on the discussion above, given a point estimate of the interference model parameters $\hat{\theta}_{\mathbf{i}}$, the symbol posterior can be written as

$$P(s_k|\mathbf{y}) = \int_{\theta_{\mathbf{i}}} P(s_k|\mathbf{y}, \theta_{\mathbf{i}})\, p\left(\theta_{\mathbf{i}}\right)\, \mathrm{d}\theta_{\mathbf{i}} \approx P\left(s_k|\mathbf{y}, \hat{\theta}_{\mathbf{i}}\right), \quad \forall k \in \mathcal{D}. \tag{6.6}$$

The MAP detection rule can then be formulated as

$$\hat{s}_k = \underset{s_k \in \mathbb{S}}{\arg\max}\, P\left(s_k|\mathbf{y}, \hat{\theta}_{\mathbf{i}}\right), \quad \forall k \in \mathcal{D} \tag{6.7}$$

where $\mathbb{S}$ is the signaling constellation. Now, given the point estimate $\hat{\theta}_{\mathbf{i}}$, this is the same MAP problem as in Chapter 5 with the true parameters replaced by the point estimate.

As discussed in Chapter 5, even with the point estimate the high-dimensional marginalization required to perform this MAP detection is intractable. Recall the OFDM system model given by

$$\mathbf{y} = \mathsf{H} \circ \mathbf{s} + \mathbf{Fg} + \mathbf{Fi} = \mathsf{H} \circ \mathbf{s} + \mathsf{I} + \mathsf{G} \tag{6.8}$$

where I have explicitly written the noise as a sum of background and impulsive noise. While, G is also an AWGN vector with independent components, this is not the case for I whose components are in fact dependent (See discussion in Section 2.1 and Section 5.4). Due to this coupling, the independent and

disjoint decoding of the OFDM subchannels as done under AWGN is no longer optimal. Using the *law of total probability*, the posterior symbol probability used in (6.7) can be written as

$$\begin{aligned} P\left(s_k|\mathbf{y};\Theta,\hat{\theta}_{\mathbf{i}}\right) &= \sum_{\mathbf{s_{\backslash k}}} P\left(\mathbf{s}|\mathbf{y};\Theta,\hat{\theta}_{\mathbf{i}}\right) \propto \sum_{\mathbf{s_{\backslash k}}} p\left(\mathbf{y}|\mathbf{s};\Theta,\hat{\theta}_{\mathbf{i}}\right) P(\mathbf{s}) \\ &= \sum_{\mathbf{s_{\backslash k}}} \int_{\mathbf{i},\mathbf{h}} p\left(\mathbf{y}|\mathbf{h},\mathbf{i},\mathbf{s};\Theta\right) p\left(\mathbf{h};\theta_{\mathbf{h}}\right) p\left(\mathbf{i};\hat{\theta}_{\mathbf{i}}\right) P(\mathbf{s}) \\ &= \sum_{\mathbf{s}_{\backslash k},\mathbf{z}} \int_{\mathbf{i},\mathbf{h}} \prod_{k=0}^{N-1} p\left(y_k|s_k,\mathbf{h},\mathbf{i};\mathcal{D},\mathcal{N},\mathcal{P}\right) \\ &\quad \times P(s_k)\, p\left(i_k|z_k,\hat{\theta}_{\mathbf{i}}\right) P\left(z_k|\hat{\theta}_{\mathbf{i}}\right) \prod_{l=0}^{L-1} p\left(h_l\right) \end{aligned} \tag{6.9}$$

where $\mathbf{s_{\backslash k}} = [s_1, \cdots, s_{k-1}, s_{k+1}, \cdots, s_N]^T$ and $\Theta\left\{\mathcal{D},\mathcal{P},\mathcal{N},\theta_{\mathbf{h}}\right\}$ are the remaining system parameters as described in Section 5.4. The coupling between the subchannels is evident in the $p\left(y_j|s_j,\mathbf{h},\mathbf{i},\mathcal{D},\mathcal{P},\mathcal{N}\right)$ term where the received signal at each tone y_j depends on the complete vectors $\mathbf{i}$ and $\mathbf{h}$ through the linearly mixed terms $\mathsf{I}_j = [\mathbf{Fi}]_j$ and $\mathsf{H}_j = [\mathbf{Fh}]_j$, respectively. This prevents the high-dimensional integrals in (6.19) from simplifying into N scalar integrals, as would happen under AWGN, and leads to an intractable marginalization.

A similar challenge arises when trying to estimate the impulsive noise parameters $\hat{\theta}_{\mathbf{i}}$. Typically, these parameters are estimated using the *empirical Bayes* framework. This amounts to *maximum likelihood* (ML) estimation of the interference prior (the impulsive noise model) parameters. Mathematically, this can be expressed as follows

$$\hat{\theta}_{\mathbf{i}} = \arg\max_{\theta_{\mathbf{i}}} p\left(\mathbf{y}|\theta_{\mathbf{i}}\right). \tag{6.10}$$

However, $p\left(\mathbf{y}|\theta_{\mathbf{i}}\right)$ requires marginalization over the transmitted symbols, the channel, and the impulsive noise. This again involves high-dimensional integrals that are intractable. In particular,

$$\begin{aligned} p\left(\mathbf{y}|\theta_{\mathbf{i}}\right) &= \sum_{\mathbf{s}} p\left(\mathbf{y}|\mathbf{s};\Theta,\theta_{\mathbf{i}}\right) P(\mathbf{s}) \\ &= \sum_{\mathbf{s}} \int_{\mathbf{i},\mathbf{h}} p\left(\mathbf{y}|\mathbf{h},\mathbf{i},\mathbf{s};\Theta\right) p\left(\mathbf{h};\theta_{\mathbf{h}}\right) p\left(\mathbf{i};\theta_{\mathbf{i}}\right) P(\mathbf{s}) \\ &= \sum_{\mathbf{s},\mathbf{z}} \int_{\mathbf{i},\mathbf{h}} \prod_{k=0}^{N-1} p\left(y_k|s_k,\mathbf{h},\mathbf{i};\mathcal{D},\mathcal{N},\mathcal{P}\right) \\ &\quad \times P(s_k)\, p\left(i_k|z_k,\theta_{\mathbf{i}}\right) P(z_k|\theta_{\mathbf{i}}) \prod_{l=0}^{L-1} p\left(h_l\right) \end{aligned} \tag{6.11}$$

which again involves the coupling term leading to the high-dimensional integral.

6.4.2 Message Passing Receivers with EM Parameter Estimation

To tackle the two challenging inference problems given in (6.9) and (6.11), I follow the same approach as in Chapter 5 and use graphical modeling with message passing to approximate these inference problems. Figure 6.1 represents the factor graph corresponding to the pdf factorization given in (6.9) and (6.11). This is similar to the factor graph of the previous chapter given in Figure 5.1. The main difference is in the presence of additional nodes representing the GM interference parameters $\boldsymbol{\pi}$ and $\boldsymbol{\gamma}$. In this section, I assume that the number of terms in the Gaussian mixture is known. Future work will address the case where this number can be estimated from the data.

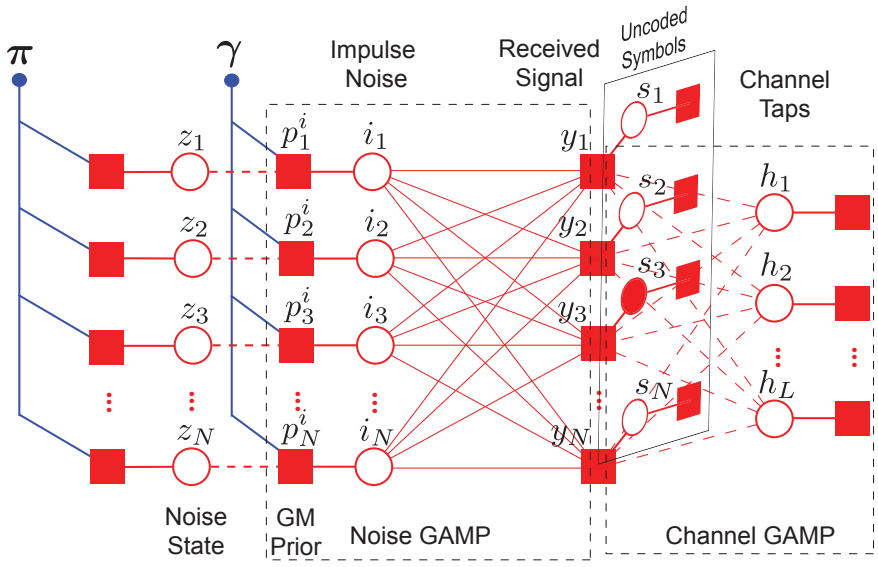


Figure 6.1: Factor graph representation of the OFDM system model given by the pdf factorization in (6.9): The blue edges and the two nodes represent the messages and the parameters of the GM model that are estimated by an EM algorithm.

For a given $\boldsymbol{\pi}$ and $\boldsymbol{\gamma}$, the messages passed between the nodes are exactly the same as those of the JCNED algorithm described in Section 5.5.3. As a result, I encourage the reader to review the corresponding section discussing the JCNED receiver before proceeding. For the purpose of this section, I assume that all the messages passed on the red-colored subgraph have been returned by the JCNED algorithm. In particular, consider the messages entering the impulse noise node i_k given by $\mathcal{N}(i_k; \hat{r}_k, \gamma^r)$ where $\hat{r}_k$ and γ^r[1] are given in steps

[1]Note that γ^r is independent of the index k due to utilizing the DFT matrix in steps (R6-7) of Table 5.1.

(R6-7) of Table 5.1. By the sum-product algorithm, we have

$$p\left(\mathbf{y}|\theta_{\mathbf{i}}\right) \approx \prod_{k=1}^{N} \mathcal{N}\left(i_k; \hat{r}_k, \gamma^r, \theta_{\mathbf{i}}\right). \tag{6.12}$$

As discussed in [79], this can be interpreted as having N observations $\{\hat{r}_k\}_{k=1}^{N}$ with noise power γ^r; the approximate system model is then given by

$$\hat{r}_k = i_k + \mathcal{N}\left(0, \gamma^r\right). \tag{6.13}$$

As a result, the ML estimation of the parameters $\hat{r}_k$ and γ^r can be approximated by

$$\hat{\theta}_{\mathbf{i}} = \arg\max_{\theta_{\mathbf{i}}} p\left(\hat{\mathbf{r}}|\theta_{\mathbf{i}}\right) \tag{6.14}$$

where $\hat{\mathbf{r}} = [\hat{r}_1, \cdots, \hat{r}_N]^T$ and $\hat{r}_k$ is given by (6.13). The solution to (6.14) can be obtained by the EM algorithm [12]. The EM algorithm, described in Section 4.6, is an iterative procedure that updates its parameters through two steps: the expectation step (E), and the maximization step (M). The derivation of the EM updates for the estimation problem in (6.14) is relatively simple and proceeds as follows: Given the current estimate of the parameters given by $\boldsymbol{\pi}^{(i)}$ and $\boldsymbol{\gamma}^{(i)}$, the value of the parameters at the next iteration is given

by

$$\begin{aligned}
\nu_{k,l} &= \frac{\pi_l^{(i)} \mathcal{N}\left(\hat{r}_k; 0, \gamma^r + \gamma_l^{(i)}\right)}{\sum_{j=0}^{K-1} \pi_j^{(i)} \mathcal{N}\left(\hat{r}_k; 0, \gamma^r + \gamma_j^{(i)}\right)} \\
N_l &= \sum_{k=1}^{N} \nu_{k,l} \\
\pi_l^{(i+1)} &= N_l / N \\
\gamma_l^{(i+1)} &= \frac{1}{N_l} \sum_{k=1}^{N} \nu_{k,l} \|\hat{r}_k\|^2 - \gamma^r
\end{aligned} \tag{6.15}$$

where K is the number of Gaussian components in the mixture. As discussed in Section 4.6, the performance of the EM algorithm depends heavily on the initial conditions. For my experiments, I choose $\boldsymbol{\pi}^{(0)}$ to be uniform and I set $\boldsymbol{\gamma}^{(0)}$ to the centers of the clusters obtained by running the *k-means algorithm* [12] on $\|\hat{r}_k\|^2$. Using this initialization, the EM algorithm produces good estimates within the first 20 iterations.

I now proceed to describe the complete Blind JCNED algorithm. Initially, I run the JCNED algorithm, described in Section 5.5.3, assuming the interference is AWGN. After convergence, I pass the resulting $\hat{\mathbf{r}}$ and γ^r to the EM algorithm described in (6.15) to produce estimates of $\boldsymbol{\pi}$ to be uniform and I set $\boldsymbol{\gamma}$. These estimates are then used as the interference model parameters for the JCNED for the next iteration. This continues until convergence or till the maximum number of iterations has been reached.

6.5 Joint Channel/Noise Estimation and Decoding with ARD Prior

In this section, I propose to extend the framework used in Chapter 5 to the case where the interference prior model is set to the *automatic relevance determination* (ARD) prior [93]. The ARD prior is a hierarchical model that promotes sparsity [93]. Typically, the inference is performed using an empirical Bayes framework called *Sparse Bayesian Learning* (SBL).

6.5.1 Sparse Bayesian Learning

I describe the model specification in sparse Bayesian learning and review the associated inference procedures proposed in [93].

6.5.1.1 System Model

SBL is commonly applied for sparse signal recovery from compressed measurements. The typical system model for compressive measurement is given by

$$\mathbf{y} = \mathbf{A}\mathbf{w} + \mathbf{n}, \tag{6.16}$$

where $\mathbf{A}$ is a known $M \times N$ measurement matrix, and $\mathbf{n}$ is a $N \times 1$ vector of *i.i.d.* noise samples drawn from $\mathcal{N}(0, \sigma^2)$. Here, the $N \times 1$ unknown vector $\mathbf{w}$ is typically high-dimensional and *sparse*, meaning that many of its elements are zeros. Typically the number of measurements M is significantly less than N.

From (6.16), the *likelihood* of the observations given the model is given

by

$$p\left(\mathbf{y}|\mathbf{w},\sigma^2\right)=\mathcal{N}\left(\mathbf{y}|\mathbf{A}\mathbf{w},\sigma^2 I\right), \tag{6.17}$$

SBL places the following *prior* over $\mathbf{w}$:

$$p\left(\mathbf{w}|\boldsymbol{\alpha}\right)=\prod_{i=1}^{d}\mathcal{N}\left(w_i|0,\alpha_i\right)=\mathcal{N}\left(\mathbf{w}|0,C\right), \tag{6.18}$$

where C is a diagonal matrix with $\boldsymbol{\alpha}$ in the diagonal. This is the *ARD* prior. Each of the N independent hyperparameters $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_N)^T$ controls the variance of its corresponding weight. This type of prior ultimately leads to a sparse model by setting many of α's to sufficiently large and associated weights to zero.

6.5.2 Inference

Given the ARD prior in (6.18) and the likelihood (4.11), the posterior over $\mathbf{w}$ from *Bayes* rule is given by:

$$p\left(\mathbf{w}|\mathbf{y},\boldsymbol{\alpha},\sigma^2\right)\propto\mathcal{N}\left(\mathbf{w}|\boldsymbol{\mu},\boldsymbol{\Sigma}\right), \tag{6.19}$$

where $\boldsymbol{\Sigma}=(\frac{1}{\sigma^2}\mathbf{A}^T\mathbf{A}+\mathbf{C})^{-1}$ and $\boldsymbol{\mu}=\frac{1}{\sigma^2}\boldsymbol{\Sigma}\mathbf{A}^T\mathbf{y}$. The posterior depends on the hyperparameters. Integrating out those hyperparameters is not analytically tractable. It is widely-used to set those to a most-probably point estimate of $\boldsymbol{\alpha}_{MP}$ and σ^2_{MP} by maximizing the *marginal likelihood*:

$$p\left(\mathbf{y}|\boldsymbol{\alpha},\sigma^2\right)=\int p\left(\mathbf{y}|\mathbf{w},\sigma^2\right)p\left(\mathbf{w}|\boldsymbol{\alpha}\right)=\mathcal{N}\left(\mathbf{y}|0,\boldsymbol{\Lambda}\right), \tag{6.20}$$

where $\boldsymbol{\Lambda}=(\sigma^2\mathbf{I}+\mathbf{A}\mathbf{C}\mathbf{A}^T)$. Although the marginal likelihood is in closed form as a Gaussian in $\mathbf{y}$; $\boldsymbol{\alpha}_{MP}$ and σ^2_{MP} cannot be expressed in closed form, and

are typically estimated by an EM algorithm. The main drawback is that computing the re-estimation functions requires inverting the posterior covariance matrix $\mathbf{\Sigma}$. This operation requires $O(N^3)$ complexity for computation and $O(N^2)$ for memory storage.

In [57], SBL was proposed as a robust approach for impulsive noise estimation and mitigation without constraining it a specific model. This SBL receiver is able to achieve significantly better performance under practical impulsive noise scenarios than most compressive sensing techniques discussed in Section 2.4. A limitation of SBL is that it depends on linear channel estimation to utilize the pilot tones which is suboptimal in impulsive noise. Furthermore, its objective is the accurate reconstruction of impulsive noise rather than minimizing the detection errors. In the following section, I propose a message passing SBL that doesn't suffer from these limitations.

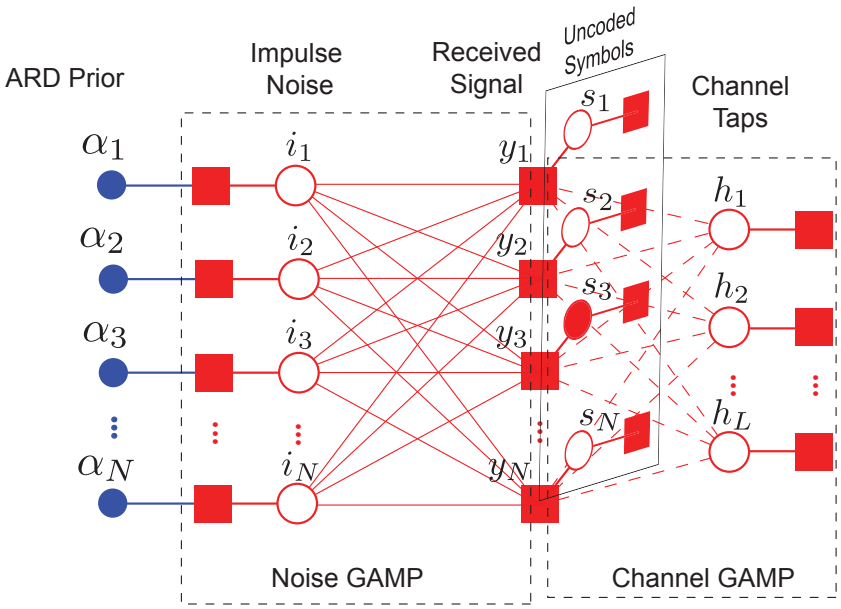


Figure 6.2: Factor graph representation of the OFDM system model when the interference model is specified by the ARD prior: the $\boldsymbol{\alpha}$ are a set of parameters that are estimated by empirical Bayes.

6.5.3 Message Passing Receivers with ARD Prior

Similarly to Blind JCNED discussed in Section 6.4, the posterior symbol probability given the received signal is given by

$$\begin{aligned} P(s_k|\mathbf{y};\Theta,\boldsymbol{\alpha}) &= \sum_{\mathbf{s_{\backslash k}}} P(\mathbf{s}|\mathbf{y};\Theta,\boldsymbol{\alpha}) \propto \sum_{\mathbf{s_{\backslash k}}} p\left(\mathbf{y}|\mathbf{s};\Theta,\boldsymbol{\alpha}\right) P(\mathbf{s}) \\ &= \sum_{\mathbf{s_{\backslash k}}} \int_{\mathbf{i},\mathbf{h}} p\left(\mathbf{y}|\mathbf{h},\mathbf{i},\mathbf{s};\Theta\right) p\left(\mathbf{h};\theta_{\mathbf{h}}\right) p\left(\mathbf{i}|\boldsymbol{\alpha}\right) P(\mathbf{s}) \\ &= \sum_{\mathbf{s}_{\backslash k},\mathbf{z}} \int_{\mathbf{i},\mathbf{h}} \prod_{k=0}^{N-1} p\left(y_k|s_k,\mathbf{h},\mathbf{i};\mathcal{D},\mathcal{N},\mathcal{P}\right) \\ &\quad \times P(s_k)\,\mathcal{N}\left(i_k;0,\alpha_k\right) \prod_{l=0}^{L-1} p\left(h_l\right) \end{aligned} \tag{6.21}$$

where I substituted the ARD prior form. The factor graph corresponding to this factorization is given in Figure 6.2. SBL inference in this context can be formulated as the ML estimation of the hyperparameters $\boldsymbol{\alpha}$ given by

$$\hat{\alpha}_k = \arg\max_{\alpha_k} p\left(\mathbf{y}|\alpha_k\right) \approx \arg\max_{\alpha_k} \mathcal{N}\left(\hat{r}_k; 0, 1/\alpha_k + \gamma^r\right) \tag{6.22}$$

where $\hat{r}_k$ and γ^r are given in steps (R6-7) of Table 5.1 and I utilized the approximation employed by GAMP and discussed in Section 5.5.2. After some arithmetic, it can be shown that the solution to the optimization problem in (6.22) is given by

$$\hat{\alpha}_k = \begin{cases} 1/(\|\hat{r}_k\|^2 - \gamma^r) & \text{if } \|\hat{r}_k\|^2 > \gamma^r \\ \infty & \text{o.w.} \end{cases} \tag{6.23}$$

where $\alpha_k = \infty$ implies that $i_k = 0$.

Now, I proceed to describe the complete message passing receiver that utilizes the ARD prior, called ARD-JCNED. Initially, I set all $1/\alpha_k$ to an average value of interference power. Then, I alternate between running JCNED of Section 5.5.3 and computing the αs.

6.6 Numerical Results

In this section, I study the performance of my proposed robust message-passing receivers by the means of Monte-Carlo simulations. In my simulations, I consider a $N = 256$ (narrowband PLC systems) subcarriers modulated by 4-QAM. I assume a dispersive Rayleigh fading channel with 5 taps. Unless stated otherwise, pilot tones are assigned uniformly while the null tones are

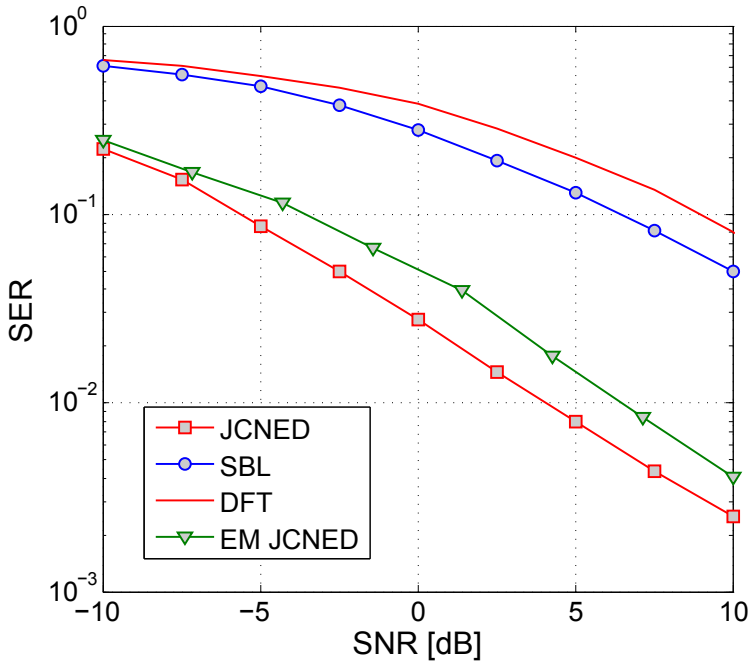


Figure 6.3: Uncoded SER of comparing Blind JCNED vs JCNED with perfect knowledge of parameters for 4-QAM and $N = 256$ subcarriers with $N_{\mathsf{n}} = 80$ null subcarriers and $N_{\mathsf{p}} = 15$ pilot subcarriers.

chosen randomly. The realizations of the impulsive noise were generated from an iid GM model having two impulsive components with powers 20dB and 30dB above the background noise occurring 7% and 3% or the time.

6.6.1 Blind JCNED vs JCNED

Figure 6.3 shows the uncoded symbol error rate (SER) comparison of my proposed Blind JCNED framework that estimates the interference parameters from the received data and the JCNED receiver from Chapter 5 which assumes knowledge of the interference parameters. The Blind JCNED suffers in the low

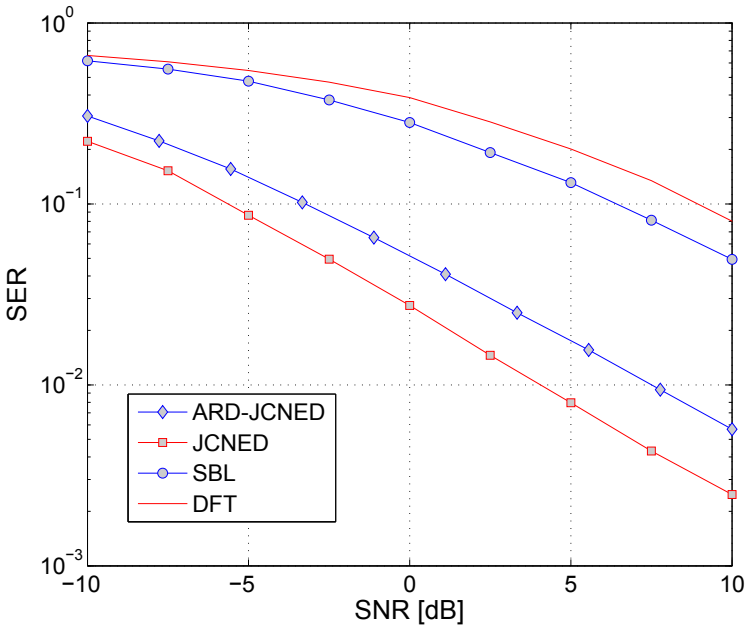


Figure 6.4: Uncoded SER of comparing ARD JCNED vs SBL receiver proposed in [57] for 4-QAM and $N = 256$ subcarriers with $N_{\mathsf{n}} = 80$ null subcarriers and $N_{\mathsf{p}} = 15$ pilot subcarriers.

SNR region due to sensitivity of signal detection to mismatch in interference model. However, at moderate to high SNRs the Blind JCNED is within 3dB from the JCNED receiver and provides gains up to 8dB over the DFT receiver.

6.6.2 ARD-JCNED vs SBL

Figure 6.4 shows the uncoded symbol error rate (SER) comparison of my proposed ARD-JCNED receiver and the SBL receiver proposed in [57]. In addition to being computationally simpler since it avoids all matrix inversions, the ARD-JCNED receiver outperforms the typical SBL implementation by around 9dB due to the fact that it jointly estimates the channel as well.

Furthermore, it is within 3dB from JCNED that assumes complete knowledge of the interference model.

6.7 Conclusion

In this chapter, I extend the graphical model framework, proposed in Chapter 5, to perform parameter estimation of the interference model. In addition, I propose replacing the interference model priors by the *automatic relevance determination* (ARD) prior that can take advantage of temporal sparsity commonly present in the interference using an empirical Bayesian framework called *sparse Bayesian learning* (SBL). Both approaches provide significant gains (around 8dB) over the DFT receiver.

Chapter 7

Conclusion

In this dissertation, I focus on designing OFDM receivers in interference limited systems and propose the following thesis statement:

In interference-limited multicarrier communication systems, accurate statistical modeling of the interference enables the design of low-complexity message passing multicarrier receivers that increase the link spectral efficiency by several bits/s/Hz, without any coordination or knowledge of the number, locations, or types of interference sources.

In the following section, I discuss how my contributions in each chapter contribute toward defending this thesis statement.

7.1 Summary

In this dissertation, I propose various statistical models of uncoordinated interference in wireless and PLC networks. Then, I leverage these models as prior knowledge in the design of OFDM receivers for interference limited environments. In particular, I propose a probabilistic graphical model framework for designing these receivers under different conditions and constraints. The specific conclusions made from each contribution are listed below:

In Chapter 3, I derive statistical-physical models for uncoordinated interference in PLC networks. In particular, I show and verify by simulation that the interference in homogeneous PLC networks follows the Middleton class-A model while the interference in dense PLC networks follows the more general Gaussian mixture model. This is inline with experimental studies reported in the literature [26, 104] and similar to models derived for wireless networks [41, 42]. In addition, I show that a cyclostationary model is appropriate for modeling the periodicity exhibited by interference in narrowband PLC and that a Gaussian HMM model captures the statistics of bursty interference present in some wireless platforms operating in the 2.4GHz ISM band. Subsequently, I use these models in the design of OFDM receivers corrupted by interference.

In Chapter 4, using a latent variable interpretation of the interference model, I propose an EM-based OFDM receiver for impulsive noise channels that is constrained to perform independent subcarrier decoding. Compared to the conventional DFT receiver, the proposed receiver provides a gain of around 6dB in the low and moderate SNR range and about 2dB in the high SNR range. To achieve this gain, the EM receiver uses two aspects of the communication system: 1) OFDM modulated signals can be approximated as being iid Gaussian in time-domain (by the central limit theorem); and 2) the knowledge of the interference model pdf and its parameters. When combined under the EM framework, these facts lead to a simple scalar LMMSE estimator in the time-domain followed by DFT detection in frequency domain. While

computationally attractive, such a disjoint decoding doesn't allow this receiver to fully exploit the OFDM's resilience to impulsive interference. Furthermore, as most of the prior work, this EM framework doesn't explicitly consider channel estimation. This can be a practical limitation as the conventional LMMSE channel estimation is highly suboptimal in non-AWGN environments.

To address the limitations of EM-based receiver, I propose in Chapter 5 a fully Bayesian inference framework to design LDPC-coded OFDM receivers in uncoordinated interference using the GM and GHMM models as priors. In particular, I propose a factor-graph-based approach to joint channel/noise-estimation-and-decoding (JCNED) of orthogonal frequency division multiplexing (OFDM) systems in interference-limited environments. My receiver merges prior knowledge of the impulsive and bursty noise models with the recently proposed "generalized approximate message passing" (GAMP) algorithm, and soft-input soft-output decoding through the sum-product framework. Unlike the prior work, I explicitly consider channel estimation in the problem formulation. Compared to the conventional DFT receiver, the proposed message-passing algorithm achieves a gain of 15dB for uncoded systems and about 13dB for LDPC-coded systems. In fact, I demonstrate that the proposed receiver is within 1dB from a lower bound on the communication performance. In addition to that, I illustrate that accurate modeling of the interference is especially important with higher order constellations such as 16QAM. Furthermore, the allocation of the pilot and null tones can have a large impact on the resulting communications performance.

In Chapter 6, I extend the graphical model framework, proposed in Chapter 5, to perform parameter estimation of the interference model. In addition, I propose replacing the interference model priors by the *automatic relevance determination* (ARD) prior that can take advantage of temporal sparsity commonly present in the interference using an empirical Bayesian framework called *sparse Bayesian learning* (SBL). Both approaches still provide significant gains (around 8dB) over the DFT receiver.

7.2 Future Research Directions

In this section, I present some promising areas for future research direction for uncoordinated interference modeling and mitigation.

1. **Temporal Modeling of Uncoordinated Interference in Wireless and Powerline Networks:** The temporal modeling of interference has for the most part been overlooked due to its difficulty. An exception is [40] where the joint statistics of uncoordinated interference in decentralized networks are shown to follow a multivariate Gaussian mixture. However, this model is too complex to be used in low-complexity receivers. As a result, there is a need for statistical-physical models that capture the temporal statistics of uncoordinated interference. Markov models are particularly appropriate since they provide a trade-off between modeling accuracy and computationally complexity.

2. **Pilot and Null Tone Allocation in Impulsive Noise Channels:**

Section 5.6.4 demonstrated that the locations of the null and pilot tones can have a significant impact on communication performance. In fact, the locations of these tones will determine the effective measurement matrix as seen from the viewpoint of impulsive interference reconstruction. This opens new interesting research directions for known tone allocation in impulsive noise channels where conventional allocation techniques for the AWGN channel fail to capture impulsive noise estimation accuracy. Furthermore, the placement of both pilot and null tones will provide a trade-off between channel estimation accuracy and impulsive noise estimation accuracy. In Section 5.5.5, I propose to use *dictionary coherence*, a metric typically used in compressive sensing, as a metric for optimizing tone allocation. However, I don't show any formal results regarding its optimality for communication performance. Furthermore, maximizing this metric is a combinatorial problem in general and thus difficult to solve in real-time. Future work can try to address these questions.

3. **Adaptive Modulation and Coding in Impulsive Noise Channels:** The proposed receiver in Section 5 and Section 6 approximates MAP inference for a specific code and bit loading scheme. However, it does not specify explicitly how such a scheme should be chosen. This problem is typically referred to as adaptive modulation and coding (AMC) [37]. In AWGN, AMC targeting either probability of error or throughput can be simply implemented independently on each subcarrier based on its SNR . This is because independent minimum distance decoding on each

subcarrier is optimal (with CSI). While it is possible to obtain the noise distribution on each carrier [36], performing AMC independently on each subcarrier will be suboptimal for either error rate or throughput. For example, while loading a high-SNR subcarrier with a larger constellation might increase the throughput for that tone, it might reduce the ability to estimate the impulsive noise thereby reducing the total throughput. Future work would target designing an AMC strategy that would optimize the total throughput or average error rate across all tones.

4. **Message-Passing Receiver Extension to Different Interference and Noise Models:** This dissertation focuses mainly on the GM and GHMM interference models. These models capture the temporal variation in noise power; however, the resulting samples are uncorrelated (GM samples are independent, HMM with zero-mean Gaussian emission probabilities [32]). As a result the spectral decomposition of these models is white. However, many communication systems exhibit interference that is spectrally shaped: the noise samples are correlated. Such a correlation may arise due to the signal propagation channel or due to the limited bandwidth of the receiver when filtering broadband or uncorrelated noise. This induces correlation in the noise and spreads impulsive events over multiple time samples. In addition, the presence of many narrowband interferers and background noise in PLC systems can further color the noise spectrum [73]. Correlation between time samples and spectral shaping of a random process are typically achieved by linear filtering

of uncorrelated samples. Stochastic process models based on this filtering approach are called ARMA models [48]. Gaussian ARMA has been used to model non-AWGN noise in narrowband PLC [70]. The proposed message-passing framework in Chapter 5 can be easily extended to accommodate such models by replacing the MC subgraph in Figure 5.1 by an appropriate inference algorithm corresponding to the utilized model.

5. **Message-Passing Receivers for Mitigating Narrowband Interferers:** While this dissertation focuses on impulsive interference in time domain, narrowband interference (NBI) arises in various powerline and wireless OFDM receivers [73]. Various compressive sensing techniques have been proposed in the literature [38]; however, these methods utilize zero padded OFDM and can not exploit the structure present in NBI [73]. The message-passing framework proposed in Chapter 5 can be utilized to design OFDM receivers that operate on cyclic prefix OFDM and, more importantly, utilize all time domain samples to estimate the NBI and cancel it out.

Bibliography

[1] Powerline related intelligent metering evolution (PRIME).

[2] F. Abdelkefi, P. Duhamel, and F. Alberge. Impulsive noise cancellation in multicarrier transmission. *IEEE Trans. Commun.*, 53(1):94–106, 2005.

[3] Fatma Abdelkefi, Pierre Duhamel, and Florence Alberge. A necessary condition on the location of pilot tones for maximizing the correction capacity in OFDM systems. *IEEE Trans. Commun.*, 55(2):356–366, 2007.

[4] R. Aggarwal, P. Schniter, and C.E. Koksal. Rate adaptation via link-layer feedback for goodput maximization over a time-varying channel. *IEEE Trans. Wireless Commun.*, 8(8):4276–4285, 2009.

[5] S. Akoum and R.W. Heath. Limited feedback for temporally correlated MIMO channels with other cell interference. *IEEE Trans. Signal Process.*, 58(10):5219–5232, 2010.

[6] N. Andreadou and F.-N. Pavlidou. Modeling the noise on the OFDM power-line communications system. *IEEE Trans. Power Del.*, 25(1):150–157, 2010.

[7] O. El Ayach, S.W. Peters, and R.W. Heath. The practical challenges of interference alignment. *IEEE Wireless Commun. Mag.*, 20(1):35–42, 2013.

[8] T.C. Banwell and S. Galli. A new approach to the modeling of the transfer function of the power line channel. *Proc. Int. Symp. Power-Line Commun. and Appl.*, pages 319–324, 2001.

[9] I. Barhumi, G. Leus, and M. Moonen. Optimal training design for MIMO OFDM systems in mobile wireless channels. *IEEE Trans. Signal Process.*, 51(6):1615–1624, 2003.

[10] D. Baron, S. Sarvotham, and R.G. Baraniuk. Bayesian compressive sensing via belief propagation. *IEEE Trans. Signal Process.*, 58(1):269–280, 2010.

[11] M. Bayati and A. Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Trans. Inf. Theory*, 57(2):764–785, 2011.

[12] Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[13] K.L. Blackard, T.S. Rappaport, and C.W. Bostian. Measurements and models of radio frequency impulsive noise for indoor wireless communications. *IEEE J. Sel. Areas Commun.*, 11(7):991–1001, 1993.

[14] T.K. Blankenship and T.S. Rappaport. Characteristics of impulsive noise in the 450-MHz band in hospitals and clinics. *IEEE Trans. Antennas Propag.*, 46(2):194–203, 1998.

[15] G. Boudreau, J. Panicker, Ning Guo, Rui Chang, Neng Wang, and S. Vrzic. Interference coordination and cancellation for 4G networks. *IEEE Commun. Mag.*, 47(4):74–81, 2009.

[16] J. Boutros and G. Caire. Iterative multiuser joint decoding: unified framework and asymptotic analysis. *IEEE Trans. Inf. Theory*, 48(7):1772–1793, 2002.

[17] G. Caire, T.Y. Al-Naffouri, and A.K. Narayanan. Impulse noise cancellation in OFDM: an application of compressed sensing. In *IEEE International Symposium on Information Theory (ISIT)*, pages 1293–1297, 2008.

[18] M.H.L. Chan and R.W. Donaldson. Amplitude, width, and interarrival distributions for noise impulses on intrabuilding power line communication networks. *IEEE Trans. Electromagn. Compat.*, 31(3):320–323, 1989.

[19] C. F. Chiasserini and R. R. Rao. Coexistence mechanisms for interference mitigation between IEEE 802.11 WLANs and Bluetooth. *Proc. IEEE International Conference on Computer Communications*, 2:590–598, 2002.

[20] T.C. Chuah. Robust iterative decoding of turbo codes in heavy-tailed noise. *IEE Proceedings on Communications*, 152(1):29–38, 2005.

[21] F. J. C. Corripio, J. A. C. Arrabal, L. D. del Rio, and J. T. E. Munoz. Analysis of the cyclic short-term variation of indoor power line channels. *IEEE J. Sel. Commun.*, 24(7):1327–1338, 2006.

[22] J.A. Cortes, L. Diez, F.J. Canete, and J.J. Sanchez-Martinez. Analysis of the indoor broadband power-line noise scenario. *IEEE Trans. Electromagn. Compat.*, 52(4):849–858, 2010.

[23] Anand Dabak, Badri Varadrajan, Il Han Kim, Marcel Nassar, and Gordon Gregg. Appendix for noise channel modeling for IEEE 1901.2, June 2011. doc: 2wg-11-0134-05-PHM5-appendix-for-noise-channel-modeling-for-ieee-1901-2.

[24] V. Degardin, M. Lienard, A. Zeddam, F. Gauthier, and P. Degauquel. Classification and characterization of impulsive noise on indoor power-line used for data communications. *IEEE Trans. Consum. Electron.*, 48(4):913–918, 2002.

[25] P.A. Delaney. Signal detection in multivariate class-A interference. *IEEE Transactions Communications*, 43(234):365–373, 1995.

[26] L. Di Bert, P. Caldera, D. Schwingshackl, and A.M. Tonello. On noise modeling for power line communications. In *Proc. Int. Symp. Power-Line Comm. and its Appl.*, pages 283–288, 2011.

[27] D. Donoho. High-dimensional data analysis: The curses and blessings of dimensionality. *American Mathematics Conference on Math Challenges of the 21st Century*, 2000.

[28] D.L. Donoho, A. Maleki, and A. Montanari. Message passing algorithms for compressed sensing. *Proc. Natl. Acad. Sci.*, 106:18914–18919, 2009.

[29] Hendrik C. Ferreira, Lutz Lampe, John Newbury, and Theo G. Swart, editors. *Power Line Communications: Theory and Applications for Narrowband and Broadbank Communications over Power Lines*. Wiley, 2010.

[30] D. Fertonani and G. Colavolpe. A robust metric for soft-output detection in the presence of class-A noise. *IEEE Transactions on Communications*, 57(1):36–40, 2009.

[31] G.J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. Technical report, Bell Labs Technical Journal, 1996.

[32] Sylvia Fruhwirth-Schnatter. *Finite Mixture and Markov Switching Models*. Springer, 2006.

[33] S. Galli, A. Scaglione, and Zhifang Wang. For the grid and through the grid: The role of power line communications in the smart grid. *Proc. IEEE*, 99(6):998–1027, 2011.

[34] C.N. Georghiades and Jae Choong Han. Sequence estimation in the presence of random parameters via the EM algorithm. *IEEE Transactions on Communications*, 45(3):300–308, 1997.

[35] D. Gesbert, M. Kountouris, R.W. Heath, Chan-Byoung Chae, and T. Salzer. Shifting the MIMO paradigm. *IEEE Signal Process. Mag.*, 24(5):36–46, 2007.

[36] M. Ghosh. Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems. *IEEE Trans. Commun.*, 44(2):145–147, 1996.

[37] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[38] A. Gomaa and N. Al-Dhahir. A sparsity-aware approach for NBI estimation in MIMO-OFDM. *IEEE Trans. Wireless Commun.*, 10(6):1854–1862, 2011.

[39] N. Gonzalez-Prelcic, C. Mosquera, N. Degara, and A. Currais. A channel model for the galician low voltage mains network. *Proc. Int. Symp. Power-Line Commun. and Appl.*, pages 365–370, 2001.

[40] K. Gulati, B. L. Evans, and S. Srikanteswara. Joint temporal statistics of interference in decentralized wireless networks. *IEEE Trans. Signal Process.*, 60(12), 2012.

[41] K. Gulati, B.L. Evans, J.G. Andrews, and K.R. Tinsley. Statistics of co-channel interference in a field of poisson and poisson-poisson clustered interferers. *IEEE Trans. Signal Process.*, 58(12):6207–6222, 2010.

[42] Kapil Gulati. *Radio Frequency Interference Modeling and Mitigation in Wireless Receivers*. PhD thesis, The University of Texas at Austin, 2011.

[43] Dongning Guo and Chih-Chun Wang. Random sparse linear systems observed via arbitrary channels: A decoupling principle. In *IEEE International Symposium on Information Theory*, pages 946–950, 2007.

[44] J. Haring and A. J. Han Vinck. OFDM transmission corrupted by impulsive noise. *Proc. Int. Symp. Powerline Communications*, 2000.

[45] Jurgen Haring. *Error Tolerant Communication over the Compound Channel*. PhD thesis, University of Duisburg-Essen, 2001.

[46] Jurgen Haring and A.J.H. Vinck. Iterative decoding of codes over complex numbers for impulsive noise channels. *IEEE Trans. Inf. Theory*, 49(5):1251–1260, 2003.

[47] Jurgen Haring and A.J.H. Vinck. Coding and signal space diversity for a class of fading and impulsive noise channels. *IEEE Trans. Inf. Theory*, 50(5):887–895, 2004.

[48] Monson Hayes. *Statistical Digital Signal Processing and Modeling*. Wiley, 1996.

[49] O.G. Hooijen. On the channel capacity of the residential power circuit used as a digital communications medium. *IEEE Commun. Lett.*, 2(10):267–268, 1998.

[50] H. Hrasnica, A. Haidine, and R. Lehnert. *Broadband Powerline Communications: Network Design*. Wiley, 2004.

[51] J. Ilow and D. Hatzinakos. Analytic alpha-stable noise modeling in a poisson field of interferers or scatterers. *IEEE Trans. Signal Process.*, 46(6):1601–1611, 1998.

[52] M. Katayama, T. Yamazato, and H. Okada. A mathematical model of noise in narrowband power line communication systems. *IEEE J. Sel. Areas in Commun.*, 24(7):1267–1276, 2006.

[53] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[54] E. Kuruoglu. *Signal processing in alpha stable environments: A least l_p approach*. PhD thesis, University of Cambridge, 1998.

[55] L. Lampe. Bursty impulse noise detection by compressed sensing. In *IEEE International Symposium on Power Line Communications and Its Applications (ISPLC)*, pages 29–34, 2011.

[56] W.R. Lauber and J.M. Bertrand. Statistics of motor vehicle ignition noise at VHF/UHF. *IEEE Trans. Electromagn. Compat.*, 41(3):257–259, 1999.

[57] J. Lin, M. Nassar, and B.L. Evans. Impulsive noise mitigation in power-line communications using sparse Bayesian learning. *IEEE J. Sel. Areas Commun.*, 31(7), 2012.

[58] Angel E. Lozano, Robert W. Heath Jr., and Jeffrey G. Andrews. Fundamental limits of cooperation. *CoRR*, abs/1204.0011, 2012.

[59] D.J.C MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[60] A.M. Maras. Locally optimum detection in moving average non-gaussian noise. *IEEE Trans. Commun.*, 36(8):907–912, 1988.

[61] Y. Matsumoto, K. Gotoh, and K. Wiklundh. Band-limitation effect on statistical properties of class-A interference. In *IEEE Proc. Int. Symp. on Electromagn. Compat.*, pages 1–5, 2008.

[62] Hidenori Matsuo, Daisuke Umehara, Makoto Kawai, and Yoshiteru Morihiro. An iterative detection for OFDM over impulsive noise channel. *Proc. Int. Symp. on Power-Line Communications and Its Applications*, 213-217, 2002.

[63] R.J. McEliece, D.J.C. MacKay, and Jung-Fu Cheng. Turbo decoding as an instance of Pearl's "belief propagation" algorithm. *IEEE J. Sel. Areas Commun.*, 16(2):140–152, 1998.

[64] G. McLachlan and D. Peel. *Finite Mixture Models*. Wiley, 2000.

[65] H. Meng, Y.L. Guan, and S. Chen. Modeling and analysis of noise effects on broadband power-line communications. *IEEE Trans. Power Del.*, 20(2):630–637, 2005.

[66] A. Mengi and A.J.H. Vinck. Successive impulsive noise suppression in OFDM. In *IEEE International Symposium on Power Line Communications and Its Applications (ISPLC)*, pages 33–37, 2010.

[67] D. Middleton. Statistical-physical models of electromagnetic interference. *IEEE Trans. Electromagn. Compat.*, 19(3):106–127, 1977.

[68] D. Middleton. Threshold detection in correlated non-Gaussian noise fields. *IEEE Trans. Inf. Theory*, 41(4):976–1000, 1995.

[69] D. Middleton. Non-Gaussian noise models in signal processing for telecommunications: new methods an results for class A and class B noise models. *IEEE Trans. Inf. Theory*, 45(4):1129–1149, 1999.

[70] M. Nassar, A. Dabak, Il Han Kim, T. Pande, and B.L. Evans. Cyclostationary noise modeling in narrowband powerline communication for smart grid applications. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3089–3092, 2012.

[71] M. Nassar and B.L. Evans. Low complexity EM-based decoding for OFDM systems with impulsive noise. In *Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 1943–1947, 2011.

[72] M. Nassar, K. Gulati, Y. Mortazavi, and B.L. Evans. Statistical modeling of asynchronous impulsive noise in powerline communication networks. *IEEE Global Commun. Conf. (GLOBECOM)*, pages 1–6, 2011.

[73] M. Nassar, Jing Lin, Y. Mortazavi, A. Dabak, Il Han Kim, and B.L. Evans. Local utility power line communications in the 3-500 khz band: Channel impairments, noise, and standards. *IEEE Signal Process. Mag.*, 29(5):116–127, 2012.

[74] M. Nassar, X.E. Lin, and B.L. Evans. Stochastic modeling of microwave oven interference in WLANs. In *IEEE International Conference on Communications (ICC)*, pages 1–6, 2011.

[75] Marcel Nassar, Kapil Gulati, Marcus DeYoung, Brian Evans, and Keith Tinsley. Mitigating near-field interference in laptop embedded wireless transceivers. *Journal of Signal Processing Systems*, 63:1–12, 2011.

[76] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. 2009.

[77] J. M. Peha. Wireless communications and coexistence for smart environments. *IEEE Pers. Commun.*, 7:66–68, 2000.

[78] R. Pighi, M. Franceschini, G. Ferrari, and R. Raheli. Fundamental performance limits of communications systems impaired by impulse noise. *IEEE Trans. Commun.*, 57(1):171–182, 2009.

[79] Sundeep Rangan. Generalized approximate message passing for estimation with random linear mixing. *ArXiv preprint abs/1010.5141*, 2010.

[80] Sundeep Rangan, Alyson K. Fletcher, Vivek K. Goyal, and Philip Schniter. Hybrid approximate message passing with applications to structured sparsity. *ArXiv preprint abs/1111.2581*, 2011.

[81] D.W. Rieken. Periodic noise in very low frequency power-line communications. In *Proc. IEEE Int. Symp. on Power Line Communications and Its Applications*, pages 295–300, 2011.

[82] B.M. Sadler. Detection in correlated impulsive noise using fourth-order cumulants. *IEEE Trans. Signal Process.*, 44(11):2793–2800, 1996.

[83] M.G. Sanchez, A.V. Alejos, and I. Cuinas. Urban wide-band measurement of the UMTS electromagnetic environment. *IEEE Trans. Veh. Technol.*, 53(4):1014–1022, 2004.

[84] M.G. Sanchez, L. de Haro, M.C. Ramon, A. Mansilla, C.M. Ortega, and D. Oliver. Impulsive noise measurements and characterization in a UHF digital TV channel. *IEEE Trans. Electromagn. Compat.*, 41(2):124–136, 1999.

[85] P. Schniter. A message-passing receiver for BICM-OFDM over unknown clustered-sparse channels. *IEEE J. Sel. Topics Signal Process.*, 5(8):1462–1474, 2011.

[86] K. Slattery and H. Skinner. *Platform Interference in Wireless Systems: Models, Measurement, and Mitigation.* Newnes (Elsevier) Publishing, 2008.

[87] E.S. Sousa. Performance of a spread spectrum packet radio network link in a Poisson field of interferers. *IEEE Trans. Inf. Theory*, 38(6):1743–1754, 1992.

[88] A. Spaulding and D. Middleton. Optimum reception in an impulsive interference environment–part I: Coherent detection. *IEEE Trans. Commun.*, 25(9):910–923, 1977.

[89] C. Studer, P. Kuppinger, G. Pope, and H. Bolcskei. Recovery of sparsely corrupted signals. *IEEE Trans. Inf. Theory*, 58(5):3115–3130, 2012.

[90] M. Tanaka. High frequency noise power spectrum, impedance and transmission loss of power line in Japan on intrabuilding power line communications. *IEEE Trans. Consum. Electron.*, 34(2):321–326, 1988.

[91] L.T. Tang, P.L. So, E. Gunawan, S. Chen, T.T. Lie, and Y.L. Guan. Characterization of power distribution lines for high-speed data transmission. In *Proc. Int. Conf. on Power System Technology*, volume 1, pages 445–450, 2000.

[92] Bo Thiesson, David Maxwell Chickering, David Heckerman, and Christopher Meek. ARMA time-series modeling with graphical models. In *Pro-*

ceedings of the 20th conference on Uncertainty in artificial intelligence, pages 552–560. AUAI Press, 2004.

[93] M. Tipping. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.

[94] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[95] Der-Feng Tseng, Y.S. Han, Wai Ho Mow, Li-Chung Chang, and A.J.H. Vinck. Robust clipping for OFDM transmissions over memoryless impulsive noise channels. *IEEE Commun. Lett.*, 16(7):1110–1113, 2012.

[96] D. Umehara, H. Yamaguchi, and Y. Morihiro. Turbo decoding in impulsive noise environment. In *Proc. IEEE Global Telecommun. Conf.*, volume 1, pages 194–198, 2004.

[97] K. Vastola. Threshold detection in narrow-band non-Gaussian noise. *IEEE Trans. Commun.*, 32(2):134–139, 1984.

[98] S. Verdu. *Multiuser Detection*. Cambridge University Press, 1998.

[99] J. Wolf. Redundancy, the discrete Fourier transform, and impulse noise cancellation. *IEEE Trans. Commun.*, 31(3):458–461, 1983.

[100] Tianyu Wu and V. K N Lau. Robust rate, power and precoder adaptation for slow fading MIMO channels with noisy limited feedback. *IEEE Trans. Wireless Commun.*, 7(6):2360–2367, 2008.

[101] X. Yang and A.P. Petropulu. Co-channel interference modeling and analysis in a Poisson field of interferers in wireless communications. *IEEE Trans. Signal Process.*, 51(1):64–76, 2003.

[102] Chi-Hsiao Yih. Iterative interference cancellation for OFDM signals with blanking nonlinearity in impulsive noise channels. *IEEE Signal Process. Lett.*, 19(3):147–150, 2012.

[103] S.V. Zhidkov. Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers. *IEEE Trans. Commun.*, 56(1):5–9, 2008.

[104] M. Zimmermann and K. Dostert. Analysis and modeling of impulsive noise in broad-band powerline communications. *IEEE Trans. Electromagn. Compat.*, 44(1):249–258, 2002.

[105] M. Zimmermann and K. Dostert. A multipath model for the powerline channel. *IEEE Trans. Commun.*, 50(4):553–559, 2002.

Vita

Marcel Nassar received the Bachelor of Engineering in Computer and Communications Engineering and a minor degree in Mathematics (with High Distinction) from the American University of Beirut, Lebanon in 2006. He became a member of the Wireless Networking and Communication Group (WNCG) in Fall 2006 upon joining The University of Texas at Austin, where he obtained his Masters of Science in 2008. Since then, he has been pursuing his PhD in Electrical Engineering at the same university. He has held four research and development internships at Intel and Texas Instruments. At Intel, he was involved in modeling interference from various non-communication sources whose emissions interfere with the WiFi band. At Texas Instruments, he worked on modeling the propagation channel and the non-stationary impulsive noise in outdoor powerline communication (PLC) systems. Furthermore, he led the team's effort on designing a class-D amplifier for Texas Instrument's PLC modem. His research interests include the application of signal processing and machine learning techniques to problems in communication systems.

Email address: mnassar@utexas.edu

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.